AURAVANA PROJECT

PROJECT FOR A COMMUNITY-TYPE SOCIETY

The Material System

AURA / SSS-MS-001 | June 2020

SOCIETAL SPECIFICATION STANDARD

auravana.org

THE AURAVANA PROJECT

SOCIETAL SPECIFICATION STANDARD

THE MATERIAL SYSTEM

Document Reference Identifier: SSS-MS-001

Date of Document Distribution: June 2020

auravana.org

To cite this publication:

- *The Material System.* (2020). Auravana Project, Societal Specification Standard, AURA/SSS-MS-001. [auravana.org]

To cite an article in this publication *(authors and article title will change)*:

- Grant, T. (2020). *The Material System Overview*. The Material System. Auravana Project, Societal Specification Standard, AURA/SSS-MS-001. [auravana.org]

ISBN: 978-1-7330651-5-3

auravana.org

GREETINGS

In an effort to provide the greatest possible clarity and value the Auravana Project has formatted the system for the proposed society (of the community-type) into a series of standard publications. Each standard is both a component of the total, unified system, as well as intended to be a basis for deep reflective consideration of one's own community, or lack thereof. These formal standards are "living" in that they are continually edited and updated as new information becomes available; the society is not ever established, its design and situational operation exists in an emergent state, for it evolves, as we evolve, necessarily for our survival and flourishing.

Together, the standards represent a replicable, scalable, and comprehensively "useful" model for the design of a society where all individual human requirements are mutually and optimally fulfilled.

The information contained within these standards represent a potential solution to the issues universally plaguing humankind, and could possibly bring about one of the greatest revolutions in living and learning in our modern time. Change on the scale that is needed can only be realized when people see and experience a better way. The purpose of the Auravana Project is to design, to create, and to sustain a more fulfilling life experience for everyone, by facilitating the realization of a better way of living.

Cooperation and learning are an integral part of what it means to be a conscious individual human. A community-type societal environment has been designed to nurture and support the understanding and experience of this valuable orientation.

The design for a community-type society provides an entirely different way of looking at the nature of life, learning, work, and human interaction. These societal standards seek to maintain an essential alignment with humankind's evolving understandings of itself, combining the world of which humans are a regenerative part, with, the optimal that can be realized for all of humanity, given what is known.

The general vision for this form of society is an urgent one considering the myriad of perceptible global societal crises. Together, we can create the next generation of regenerative and fulfilling living environments. Together, we can create a global societal-level community.

THE UNIFIED SOCIETAL SYSTEM: MATERIAL SPECIFICATION STANDARD

This publication is one of six representing the proposed standard operation of a type of society given the category name, 'community' (a community-type society). This document is a specification standard for a material system.

Every society is composed of a set of core systems. Different types of societies have different internal compositions of these systems. The composition of these systems determines the type of society. The type of society described by the Auravana Project societal standard is a, community-type society. The standard is a composition of sub-system standards. The Auravana societal standard may be used to construct and duplicate community at the global level.

For any given society, there are four primary societal sub-systems. Each of these sub-systems can be specified and standardized (described and explained); each sub-system is a standard within a whole societal specification standard. The first four primary standards of the six total standards are: a Social System; a Decision System; a Material System; and a Lifestyle System. Each standard is given the name of its information system. The fifth publication is a Project Plan, and the sixth is an Overview of the whole societal system. Together, these standards are used to classify information about society, identify current and potential configurations, and operate an actual configuration.

- **This societal specification standard is the Material System for a community-type societal system.**
- **There are more figures (and tables) associated with this standard than are identified in this document; those figures that could not fit are freely available through auravana.org, in full size, and if applicable, color.**
    - *Figures and tables on the website are named according to their placement in the standard.*

Contents

List of figures

This is the list of figures within this document.
There are more figures associated with this standard than are identified in this document; those figures that could not fit are freely available through auravana.org, in full size, and if applicable, color.

List of tables

This is the list of tables within this document.
There are more tables associated with this standard than are identified in this document; those tables that could not fit are freely available via the project's website.

Document revision history

A.k.a., Version history, change log.

This document is updated as new information becomes available.
The following information is used to control and track modifications (transformations, changes) to this document.

<table>
<tr><th>VERSION</th><th>REVISION DATE</th><th>SECTIONS</th><th colspan="2">SUMMARY (DESCRIPTION)</th></tr>
<tr><td>001</td><td>June 2020</td><td>n/a</td><td colspan="2">This is the first version of the unified release of the societal standard for a community-type society. This is the first version of the material system.

Note: The reader should understand that this document contains a high-level of conceptual linguistic detail, the reader should understand that this document is one of six total documents that together provide a complete explanation of the proposed societal system. In order to visualize the complete realization of the whole societal system, its concepts and objects, and their interrelationships, must be modeled and reasoned.

Note: All figures associated with this standard, many of which are not published herein, are all available via the project's website. It is not possible to publish via this page medium all figures related to this standard.</td></tr>
<tr><th colspan="3">GENERATION ON</th><th>NAME</th><th>CONTACT DETAIL</th></tr>
<tr><td colspan="3">June 2020</td><td>Travis A. Grant</td><td>trvsgrant@gmail.com</td></tr>
</table>

The Material System Overview

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

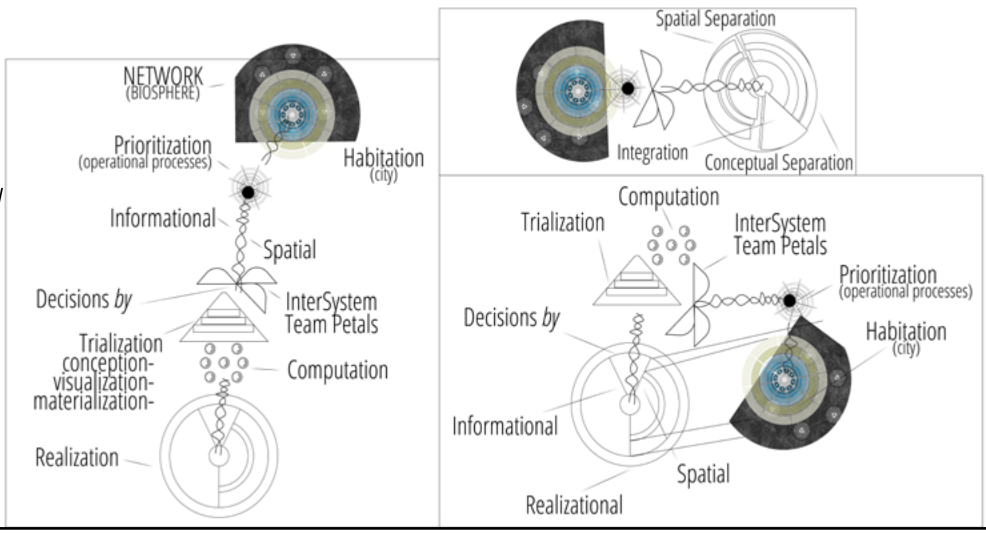
Keywords: material system, physical system, planetary environment, physical environment, material environment, ecological environment, object system, shape system, material control system, material coordination system, material configuration system, physical society

Abstract

This publication is the Material System for a community-type society. A material system describes the organized structuring of a material environment; the material structuring of community. This material system standard identifies the structures, technologies, and other processes constructed and operated in a material environment, and into a planetary ecology. A material system encodes and expresses our resolved decisions. When a decision resolves into action, that action is specified to occur in the material system. Here, behavior influences the environment, and in turn, the environment influences behavior. The coherent integration and open visualization of the material systems is important if creations are to maintain the highest level of fulfillment for all individuals. This standard represents the encoding of decisions into an environment forming lifestyles within a habitat service system. The visualization and simulation of humanity's connected material integrations is essential for maintaining a set of complex, fulfillment-oriented material constructions. As such, the material system details what has been, what is, and what could be constructed [from our information model] into our environment. This specification depicts, through language and symbols, visualization, and simulation, a material environment consisting of a planetary ecology and embedded network of integrated city systems. For anything that is to be constructed in the material system, there is a written part, a drawing part, and a simulation part, which is also how the material system is sub-divided.

Graphical Abstract

Figure 1. *Depiction of several views on the realization of a community-type societal information system that cooperatively organizes a mutual human prioritized informational and spatial service system that materializes in a physical/ material environment, primarily in the form of a network of integrated city systems, that collect and serve information and material requirements for a population of humans.*



1 Introduction to the material standard

The visualization and simulation of the interconnected material reality is essential for maintaining a set of complex material constructions designed to remain in alignment with the regeneration of our highest potential state of fulfillment (i.e., our highest potentially expressed fulfillment or HPEF). In a material context, our HPEF is determined by how efficiently we re-configure the resources in our environment into services which are effective for fulfilling human needs and maintaining individuals' access to opportunities for discovery and growth.

The Material System describes what has been, what is, and what could be constructed [from the society's information model] into its material environment, and therein, a global human habitat. In other words, it is a description of what has been materialized (past material state), could be materialized (future potential material state), and what is, materializing (current operating state). It depicts the selectively materialized expression of the society's information model as well as all probable alternatives. Essentially, this is a standard for the material domain of a society. Simply, it contains all information about the material nature of society; it describes the part of the community [information] system that is, was, or could be operational at the material level. This standard addresses the materially constructed system through knowledge and tools into processes and services that combine to form technologies which function to provide resource flows and material transformations for human fulfillment.

This standard accounts for the localized placement of all material resources within the biosphere and local habitat service systems. It accounts for not only resources, but also for the material reconstruction of the common environment through integrated access to common [heritage] resources via a unified information model, one phase of which represents the product of the interaction of the other three systems, the materialization of a service system (i.e., the after decisioning comes material reconfiguration, which humans then live within for some duration of time).

This standard depicts through written word, visualization, and simulation the materialization of an integrated habitat service system, which is more commonly referred to as a total city system network. The habitat service system is the material, technical system which facilitates the fulfillment of identified human needs. The material system is a planned. The cities herein are connected and integrated into a network of cities forming a complex human community contributing to and utilizing a unified information model and a global material access system.

NOTE: *A 'material specification' is the expression of a set of materializing relationships.*

Humans experiences their constructions, and so, in community, humans socially organize and plan for their constructions. This is accomplished through the social collection of information, which is feed into an shared decision space, which is transparently resolved into a re-configuration of the common material reality. Generally, that re-configuration takes the form of a circular walking-garden city system composed into a distributed network of cities operating together and based upon the same information model.

The integrated city system is a controlled service space where our built world mimics and harmonizes with the regenerability principles of the natural world. If we flow with natural principles we can even amplify what we are capable of in nature; we can get even better at it, and do it in a way that keeps us harmonious with the natural world, so that we aren't fighting the flow. And herein, we structure our lives and environment so that we naturally do things that are fulfilling

In community, our evolutionary impulse gives rise to the dialectic of "progress" expressed through continuous improvement within an ecological environment, rather than continuous expansion of a made up number that a bunch of people are telling themselves a story about (i.e., "GDP"). Community accounts for that which exists, and by letting go of [artificial] narratives it can optimize therein. In the Community, technology's ability to generate abundance is fully utilized without the restrictions of an economic system that values, and thus, manufactures scarcity. In community, the capacity to automate rote tasks is fully utilized, freeing humans from tedious jobs in an economic system where they are no longer needed. Technology is embraced, not as a solution in itself, but as an extension of our abilities and power, which will be as constructive as our value systems is inclusive.

1.1 Material construction

The Material Specification represents the convergence of information and matter into physical construction. In order to socially construct something into the material environment of the community [at least] three principal elements are required: a written explanation; line drawings; instrumentation and measurement. Each of these three elements is essentially a different viewport (i.e., a different window) into the same information model. We are capable of expressing our perceptions of reality via multiple mediums, such as written language, visualizations, and through ordering. When this information is combined into a single package, which can be understood by a receiving entity, then the systems design is replicable, and can be duplicated given the availability of resources.

When we encode our concepts into material structure they begin to take up space around us, which become the very constructions we live in and spend most of our lives around. From this perspective, the material specification provides descriptive reasoning for why we have constructed that which we experience as our constructions in an environment, and how to reconstruct our constructions given what is known and what is available.

In order for something to be materially constructed in community we have to know:

- Why we are constructing.
- What we are constructing.
- How we are constructing.
- The alternative ways and configurations we could be constructing.
- And, how to replicate our current and past constructions.

In other words, the written part is our description and logical reasoning for the system as it could be, as well as the system as it presently exists in its current state of operation. Importantly, it includes instructions on how to construct different versions of the city. It also describes to the constructor/builder how to build the material design for the community.

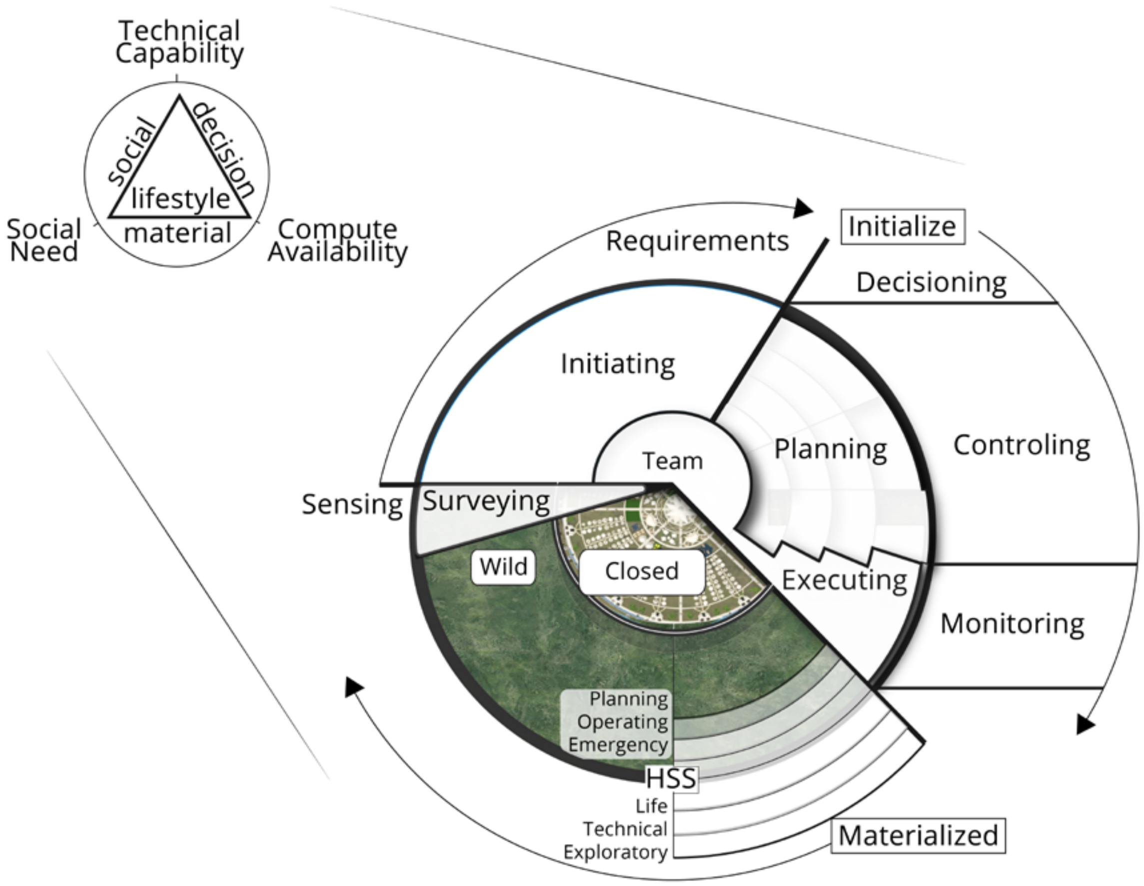
2 The material specification standard sub-composition

The material system subcomposition:

1. The written documentation part.
2. The architectural CAD- and BIM-based drawings for the integrated city system.
3. Database of materials and their properties, and technologies
4. The 3D visually modeled and simulated representation of the integrated city system.
5. Integration of the 3D representation into a gaming engine for virtually simulating all technical operational aspects of the community.

A material system describes the material blocks/ patterns that we have to work with and the optimal configuration of those patterns to sustain and evolve our fulfillment.

Figure 2. *The real-world community model executing [by means of teams and objects] the realization of a network of habitat services systems within a larger planetary ecology where humans have needs (real world requirements) that can be met with some level of knowledge and understanding.*



Material system documentation shall indicate, at a minimum, the following:

1. A description of the system functions, or a functional diagram.
2. Specifications of systems and their location (if available)
3. Type of materials
4. Type of technologies and requirements for their installation

In the sense that any given material environment can be interfaced with, there are seven primary relationships:

1. Identification: concepts & naming [conceptualizing and naming]
2. Location[ing]: positioning between objects.
3. Design[ing]: construction of an object.
4. Services[ing]: constructing the motion of multiple objects to serve a function.
5. Structure[ing]: the integration of multiple objects in motion.
6. Account[ing for] Materials: the composition of any given object.
7. Account[ing for] Technological modules (a.k.a., technological infrastructural modules): the construction and integration of multiple objects to serve a usage.
8. Account[ing for] Human requirements: the needs and preferences of the human users.

In the sense that any given material environment can have any of four primary gestalts:

1. Structures (objects on land or crafts in mediums)
2. Subjects (people or people-like organisms)
3. Energetics (motion, electromagnetics)
4. Terrain (planets and human re-contoured land)
5. Devices (functional objects in structures or on land for specific temporary and/or mobile use; a.k.a., tool, non-structure usable item; a consumable may, or may not, be considered a device)

2.1 Material specification components

There are four principal parts to the specification for the materialization of the Community.

1. Specifications - the written documentation part. A specification set may also include the drawings for the set.
2. Drawings - the graphical presentation of that which is to be constructed. Drawings are intended to depict the general configuration and layout of a design, including its size, shape, and dimensions. It informs the constructing entity of the quantities of materials needed, their placement, and their general relationship to each other. Although drawings may contain all the information about a structure that can be presented graphically, they nevertheless omit information that the contractor must have, but which is not adaptable to graphic presentation. Information in this category includes quality-related criteria for materials, specified standards of workmanship, prescribed construction methods, etc. There should be no discrepancies between drawings and written specifications.

2.2 What is a master specification?

A "master specification" is a template document that must be used and/or edited to suit a specific project. In other words, a master specification contains sufficient information that it can be extracted and/or edited for specific projects. Master specifications are also referred to when modifications are implemented to fit particular conditions of a given job or new specifications are incorporated. The Auravana Project's societal specification standard is a master specification.

In concern to a master construction specification, for example, the master may contain a list of index numbers, characteristics, specifications, units of measure, and additional information that is to be used for specific material projects.

2.3 Specifications

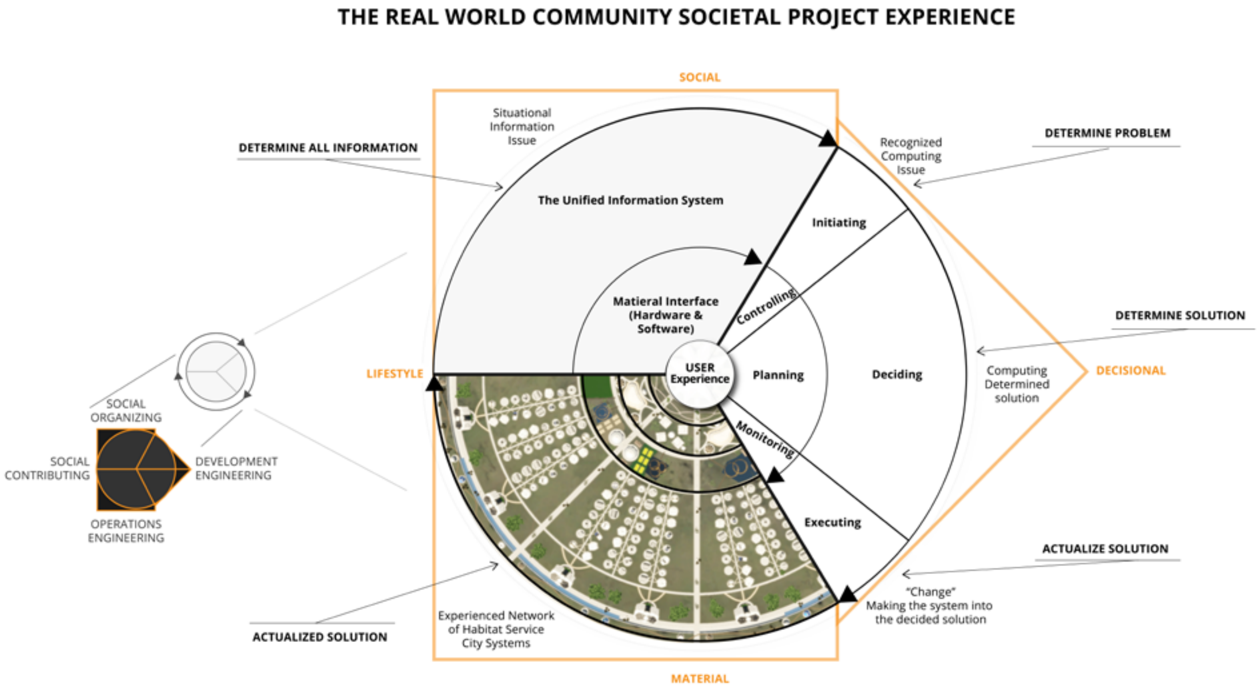
The word specification merely refers to the act of "to state explicitly or in detail" or "to be specific". There are many different types of specification. Sometimes, a "specification" is a written technical descriptions of a design, which may be contrasted with a "drawing", which is a visual depiction of a design. However, here, the total description of a design is called a "specification". Anything that is to be constructed must involve a specification. If it is needed in order to understand, construct, operate, or take down, then it is a component of a specification. A specification may also refer to a type of technical standard. In a way, these specifications represent the technical standard for community.

Note: From the commercial perspective, specifications are "that portion of the Contract Documents consisting of the written requirements for materials, equipment, systems, standards and workmanship for the Work, and performance of related services."

All specifications for the material system, including the material system itself, involve written language, symbols, drawings, and simulations. These are separated into "parts" of a specification. For anything that is to be constructed in the material system, there is a written part, a drawings part, and a simulation part, which is also how the materials system specification is itself divided. Symbols are likely to be used throughout. Here we may refer to the part of the specification which is written as "the written part of the specification", or "the written specification". The written part uses verbal language: the language of reasoning (verbal reasoning) and science (scientific evidence) to describe why the system is so constructed; and, technical language describing materials, equipment, systems standards, workmanship for the work, performance measures, and performance of related services; and engineering language to describe how it is so constructed, including composition, creation, assembly, and disassembly (as well as [re]-cycling). Visual specifications include those which are represented as [technical] drawings (i.e., drawing specifications or the drawing part) and simulations (i.e., simulation specifications or the simulation part). A technical drawing precisely and visually communicates how something functions or is to be [de-]constructed. Technical drawings are understood to have one intended meaning (i.e., they are not interpretable in more than one way) -- they use visual language to ensure they are not ambiguous and relatively easy to understand. Drawings are made according to a set of conventions, which include particular views (floor plan, section etc.), sheet sizes, units of measurement and scales, annotation and cross referencing. Herein, all architectural drawings, mechanical or other sketches, and CAD drawing applications are considered [technical] drawings. A simulation, however, is the imitation of the operation of a real-world process or system over time. The act of simulating something first requires that a model be developed; this model represents the key characteristics or behaviors/functions of the selected physical or abstract system or process. The model represents the system itself, whereas the simulation represents the operation of the system over time. Any object (i.e., "material thing") which is to be, or has been, constructed is described within a specification that includes all three parts.

Using the word "specification" without additional information to what kind of specification you refer to is confusing and considered bad practice within systems engineering. Hence, it is important to state that when we refer to the "design specifications", what we are referring to is the comprehensive specification for the total community system, which may be called the "Community System Specification". It is the core/kernel specification.

Figure 3. *This is a project to build a unified type of society where the total environment is recognized as necessary for computing a re-organization of the material environment where humans exist and persist in accordance with their behaviors, their creations, and the larger cosmic dimensional sphere. There is a material existence to human consciousness that can be physically interfaced with through human behavior (or, more precisely, an individualized consciousness behaving as a human among others similarly behaving. It is possible to orient and re-orient the materially structured environment where all humans behave together. It is possible to decide and design together a global platform composed of a network of integrated city systems where human individuals are fulfilled through the realization of a specific configuration of a socio-technical environment (i.e., a specific configuration of a material, physical environment; a specific societal object configuration environment). By planning these societal environments based upon the integration of a total information environment, it is possible to optimize material creation (Read: the material dimension of the human experience) for all of humankind and the ecological system upon which it exists.*



Specifications are mean to integrate and connect with one another. Specifications are meant to be operated and then retired or updated. The architectural construction industry organizes its specifications into divisions and levels.

> **QUESTION:** *A specification is a formalized design. Hence, the word "design" may be considered superfluous before the word specification.*

2.4 Individual technical product sheets

See addendum for individual technical product sheets, which include:

1. An individual product specification sheet.
2. A product sub-system operating parameters.

2.5 Material processes

A.k.a., Material dynamics.

There are many examples of material dynamics, including, but not limited to:

1. Biospheric/ecological dynamics
2. Water & atmospheric dynamics
3. Energy dynamics
4. Chemical dynamics
5. Structural dynamics

All systems (power systems) have resource [depletion] impacts and environmental impacts.

2.6 Material objects

Material components are the building "blocks" for creating a material system. No two or more objects can occupy the same space (spatial scarcity) at the same time (temporal scarcity).

2.7 Space control through space-time separation

In community, the population plans for well-being. A population create a setting that is conducive to better flow and better use. Here, spaces with different lifestyle-functions that could conflict are separated by time and/or space. For example, space for noisy social interaction is separated from quiet space. Not only does this separation facilitate natural movement between spaces with different lifestyle- and system-oriented functions, but it reduces disturbance for those using particular spaces with set functions.

2.7.1 Material Optimization

Human constructions become a part of the human environment. A body naturally become adapted (e.g., optimized) to its given environment. When individuals optimize the environment for our fulfillment we facilitate the experience of optimum fulfillment. It is wise to create material structures that facilitate the alignment of our behaviors with our explicit desire (i.e., intention) for purpose, potential, and play.

2.8 Economic planning and habitat elaboration

An economic plan requires an elaborate description of each habitat service subsystem. These descriptions are required in order to develop a complete input-output matrix for a habitat economic service system.

A City System for a Community-Type Society

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

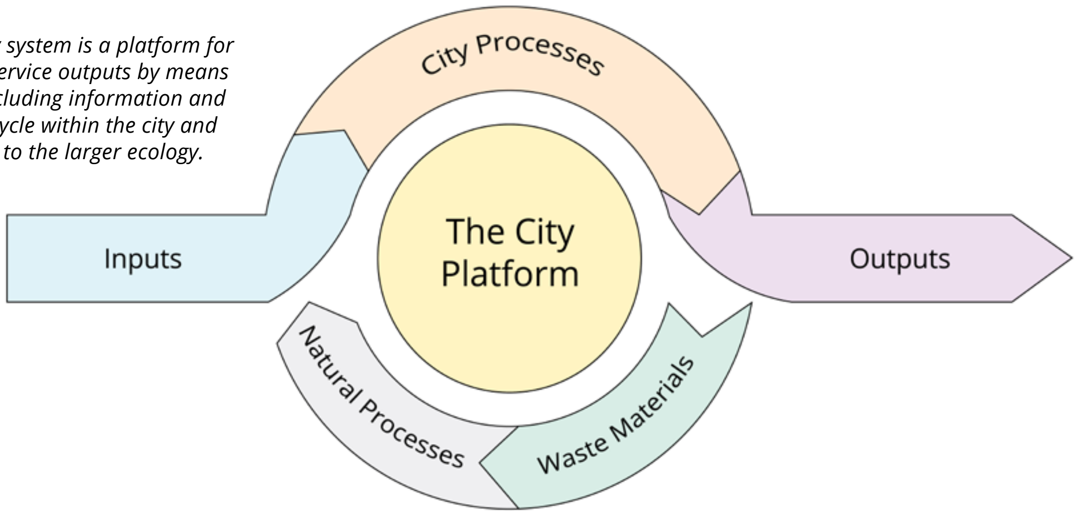
Keywords: city, city system, habitat service system, life radius, village system, cybernetic city, cybernetic habitat

Abstract

A city in a community-type society is an integrated, total living environmental service system. Another name for a city is a habitat service system. Cities in community are the result of viewing society first and foremost as an information system, from which cities are designed, developed, operated and adapted. A city represents an individual's generalized life radius and the localization of services. Community-cities are humanity's primary habitat within the larger planetary ecology. These cities form a unified cooperative network for the sharing of information and resources to maximize global fulfillment of access.

Graphical Abstract

Figure 4. *A city system is a platform for the delivery of service outputs by means of processes, including information and resources that cycle within the city and through the city to the larger ecology.*



1 Introduction

A.k.a., The habitat service system.

The technical term for a 'city' is a 'habitat service system' - a service system integrated into a larger ecological/ planetary habitat. The city is a central concept for human population-scale living. Every city is a continuous process of material creation, usage, and re-configuration. Therein, the rise of city networks and the movement of humanity into a predominantly engineered environment corresponds to a broader process of change during the "anthropocene" (the age of humans).

TERMINOLOGY: *Habitation (noun) the state or process of living in a particular place; a dwelling or city in which to live.*

In common discourse, the word "city" carries several characteristics, each portraying a different perspective on life in a materialized socio-economic system. The term, "city", is characterized differently depending upon societal perspective. From the perspective of population scale, people come together to form service hubs which are often given different names depending on the specific population size of the service hub. The population scale is generally something equivalent to: tribe [smallest population] hamlet > village > town > suburb > city > megacity (metropolis or megalopolis) [largest population]. In community, however, there is a global, unified habitat service system [network] in which there are many interconnected smaller habitat service systems known commonly as 'cities', each of which has been designed with a carrying capacity set by its [designed] configuration. To some extent, these other terms (e.g., village, town, etc.) for what is essentially a service system hub integrated into a habitat, divert attention from that fact. In other words, because of the language around how humanity is fulfilled, people don't generally think about a "village" or "city" in terms of what it actually is; which is, the materialization of socio-economic information into the material environment in the form of a habitat service system that exist to service human fulfillment requirements.

From a 21st century perspective on resource sustainability, a city is a place that requires that the majority of resources and goods come from outside of its boundary (i.e., it is, fundamentally, an unsustainable structure). Conversely, in community, the habitat service system must by definition be sustainable. It must be sustainable, because the term itself indicates the presence of a service system within a habitat, and for a habitat to sustain the population it must not overshoot carrying capacity. A habitat is a place where material needs are regeneratively fulfilled. If a habitat service system cannot sustain itself with local resources or sustainable access to the global network of resources, then it will assuredly begin to degrade and should not be considered a habitat service system. In community, the term 'habitat service system' is just another word for a "city", but the city is sustainable.

From the perspective of authority, a city is a governed area with a leader (or leadership) of some kind. Therein, a city as a space of government and authority, the territorialization of government through the structuring of social networks based upon power-over-others relationships. The city is a place where discipline and subordination is imposed. In contrast, in community, there is no authority who governs or otherwise leads everyone, or anyone. Instead, there is a common information model, an accountable decision process, and a contribution structure that facilitates participation in ones own life and the lives of all others.

In many jurisdictions (i.e., "States") cities are legally defined by their status of incorporation. City status is thought to be a natural progression to further raise the competing population's economic profile. A city, like everything else in the market, is a product to be bought and sold. The most well-known city products are Paris, Beijing, Moscow, London, and New York. Aside from marketing to get people to understand and participate in community, marketing is not a part of the socio-economic process for any city within the community network.

A city could be viewed as a piloted system, and equated with the modular nature of a space station. Cities could be designed like modular space stations, highly flexible and highly sustainable. A space station, like a habitat service system, has modular life (and other) support functions. Thus, in a more abstract sense, the core platform of a habitat service system could possibly be viewed as a "spacecraft bus", which is a major part of the structural subsystem of a spacecraft. It provides a place to attach components internally and externally, and to house modules. The bus also establishes the basic geometry of the spacecraft, and it provides the attachment points for appendages such as conduits (e.g., booms), communications elements (e.g., antennas), and sensors. Similarly, a city is a platform for the service fulfillment of humankind and its subsystems provide a variety of required functionality. However, a habitat service system is unlike a space station in that it integrates into a larger ecological environment, whereas a space station is designed to isolate its occupants for their protection. A city ought to connect and facilitate the adaptation of a life form to its environment. Also, it may provide protection, but if it isolates its users/ occupants from the larger habitat, then its users are likely to become disoriented from their highest potential trajectory in life. In reality, a service system (a city) must allow for the thoughtful flow energies with the larger ecological habitat if it is to flourish.

A city is where many people gather and live. A city is more than just an architectural expression -- it involves more service functions than just providing architecture for a population.

As a complex system, a city can express the following properties:

1. A city is a group of people and a number of permanent structures within a geographic area. In the market-State, said organization exists to facilitate social control and the trade of goods and services among residents and with the outside world. In community, said organization exists to facilitate so organized to facilitate human fulfillment.
2. A city is a center of, and for, human activity.
3. A city is a spatial pattern of human activities and their materializations at a certain point in time (note that this is often the definition for the term 'urban').
4. A city is an environment (space) configured through flows (exchanges, networks) of communication and the transport of matter and energy.
5. A city is an environment built for a social population.
6. A city is a network of people, information, and material flows designed to fulfill the requirements of a population. A city is a complex interconnection of people, information, and material flows. Therein, a city is a material system designed through the information system of a social population.
7. A 'city' is an information and spatial service system.
8. A city is an information and material processing system/platform that has been engineered for a social population.
9. A city is a socio-technically built environment that has been engineered to socio-technical specification standards. A city as a built information medium.
10. A city is the materialization of a population's socio-economic [information] system.
11. A city contains subsystems that provide structural and functional meaning, interconnected and enabled by networks of people, information, and material flow, as well as characterized by social, decisional, material, and lifestyle aspects.
12. A city is a habitat service system to fulfill human requirements.
13. A city is a multi-physic object, characterized by a dynamic interrelationship of energy, materials, information and human activities and behavior (Stepandic, 2019). In this sense, a city could be considered a "living" object of human, informational, and material interrelationship.
14. A city is an evolving, living system in complex interaction with its population, its artificially materialized environment and its natural physical environment.
15. A city is an environmentally controlled area.
16. A city is a living complex of geometrical and topological objects, limited in its material (artificial and natural) environment.
17. A city is sustained (operated) and lived in by its population, constrained by decisioning and physical laws.
18. Cities are hubs of people, supported by infrastructure. Cities are where people come together to experience the benefits of being together in close proximity - to access to services and opportunities.
19. A city is materially encoded intelligence (or, lack thereof).
20. A city is a place where people live and work and play together. A city is a live-work (or, live-work-play) environment.
21. A city is an environment that optimizes (or, should optimize) for the quality of life of the beings in that environment.

A city is a place where people live and work together under the same infrastructure:

1. City infrastructures are an enabler of collective living and working.
2. Infrastructures affect the "life" and behavior of a city.
3. Infrastructures are affected by the "life" and behaviors of a city.

A city is a place where people live within the same [type of] society:

1. A city is [the result of] a social system.
2. A city is [the result of] a decision system.
3. A city is [the result of] a material system.
4. A city is [the result of] a lifestyle system.

A city is a system with:

1. Feedback loops (inside and across sub-systems)
2. Complexity (interrelatedness)
3. Path dependency
   A. Socio-systems
   B. Technical-systems

> **INSIGHT:** *In a city, people are likely to connect only if a city's human-scale geometry creates shared spaces with the right complexity.*

1.1 Types of cities

A.k.a., Cities in community and other types of society.

In any society, cities are the product of that societies societal information set. In other words, cities arise and are reconfigured based upon a given society's information system, whether that system is made explicit or not. Different types of societies create different types of material environments, and therein, different types of cities:

1. Community-cities (a.k.a., integrated city systems, total city systems, etc.) - cities developed in community.
2. Market-State cities (a.k.a., market-city system, State-city system, smart city system, etc.) - cities developed for the market-State.

Fundamentally, specific organizations of cities play a key role in the systemic transition towards a more sustainable way of living.

1.1.1 Community-type cities

A.k.a., Community-cities, cities in community, community-based cities, community-type cities, integrated city systems, total city systems.

The term, 'community-city', is referential of the 'city' emerging as the material expression of a community-type society. Therein, 'community' is the term giving to the living system (i.e., society) as a whole, and the 'city' is its material expression. A community-city is a material ecosystem designed to facilitate the experience of a greater sense of connection and integration within each individual. Community-based cities are hubs for the sharing of access. Cities in community are complex socio-technical environments where access and services are available for free. Most cities in community are largely autonomous; however, some populations have chosen to live in low technologically developed cities where there is little application of autonomous systems. These low-tech city systems are often significantly reduced in population size and carry capacity compared to their high-tech alternatives. Effectively, community-type cities are integrated socio-technical service systems designed to fulfill the needs, wants and preferences of a population of human beings in a regenerative and emergent manner.

A community-city is an integrated (total) socio-technological service system for providing material need fulfillment to a network of humans in a sustainable manner at pre-planned population scales. Therein, a city is the materialized reflection of socio-decisioning and lifestyle design.

1.1.2 Market-State cities

A.k.a., Industrial cities, modern cities, early 21st century cities.

All known cities to date are market-State cities. A State-city or city-State refers to a population center for social interaction within a territorial authority, the "jurisdiction" (i.e., a government body). State-based cities are governmental controlled environment (i.e., population-controlled environments). The population of a city-State is often called its "citizenry" -- the population are "citizens" of a particular jurisdiction. Here, a city is a center for national/local government. The state controls (or, restrains) the behavior of citizens in a city through laws and law enforcement (i.e., the monopolization of force/violence).

A market-city or city-market refers to a city designed to accommodate the market. In the market, a city is a center for trade and financial investment. Market-based cities are designed around ownership and businesses. Individuals in a market-city are generally referred to as owners. A market-State city is necessarily divided up into plots of ownership, often, starting with that which is owned by the local governing authority. In fact, nearly all cities in early 21st century society are entirely owned by their local governing authority, which rents the land to secondary owners by means of taxation. Hence, market-cities are actually city governmental markets. Thus, a city is a conglomeration of people and buildings clustered together to serve as a center of politics, culture, and economics. Therein, a city is an environment in which people compete for resources.

TERMINOLOGY: '*Urbanization' is the growth and diffusion of city landscapes and urban lifestyles.*

Cities within the market-State are developed around entirely different purposes than those developed for/in community. Early 21st century cities are a central hubs for competition, and the vast majority of them are designed primarily for automobile access. Modern cities are full of passive commercial attractions, and capital cities are full of embassies.

In the market-State, market demand is closely related to the size of any given economy, population, and income level of a city. Therein, different aspects of the market and State determine the spatial distribution, size, direction, and model of the production and circulation of goods and services within and between cities. (Ni, 2007)

Market-State cities in the early 21st century are generally designed along five business lines (*Systems of Cities*, 2009):

1. City management, governance, and finance.
2. Economic growth.
3. City planning, land, and housing.
4. Urban environment and climate change.
5. Handling urban poverty.

Note: These business items set out the objectives and benchmarks for financing and policy advice.

The majority of data available on land and resource requirements to support a market-State city in the early 21st century is mostly irrelevant to the design of a city as a habitat service system in community. For example, the amount of food volume necessary to support the population is different since a community's service system accounts for nutritional density, which is something modern socio-economic service entities do not account for in their data. Further, there is no financial system in a community-city, which makes data tainted by financial bias somewhat useless.

> **NOTE:** *Market-State cities manage the poor instead of creating environments that don't produce poverty.*

In the early 21st century, some market-State cities are centuries old, having been built up and outwards for thousands of years. The lives of the current inhabitants of these cities are significantly shaped by centuries old structures and layouts that are creations of those with knowingly outdated understandings and value sets. In other words, the beliefs and values of those long dead are still affecting the day-to-day lives, behaviors and lifestyles, of city residents in the early 21st century.

Living systems survive, connect with their environment, and reproduce themselves. The biophysical facts of life then set up the conditions for individual survival and species survival. The city is just a natural and inevitable outcome of human behaviors that have resulted from human evolution. But, it has also become a life condition itself that directly impacts well-being. Cities in the early 21st century (and their suburbs) represent the structural encoding of a value orientation away from one of a resilient living system. To flourish, humanity must redesign its city habitats to more greatly encode and restore a life-fulfilling value set.

It would be far easier and would require less energy to build new, efficient cities than to attempt to update and solve the problems of the old ones. (Circular cities, 2020) The question then arises, what would be done with old cities in a global community-type society? Most of the old cities would be leveled and mined for their resources. They are too inefficient to maintain. Some of the cities would be set aside as museum cities. (FAQ, 2020)

> **INSIGHT:** *In making the city, we make the world more after our "hearts desire", but in making that city we also make our future selves.*

1.1.3 Smart cities

> *A.k.a., Smart urbanism, electronic government, electronic infrastructure, urban operating system.*

Smart cities are, generally, more technologically advanced cities in the market-State. A smart city is the conception of a city that uses technology to enhance governance, planning, management, and livability of a city by gathering and processing real-world, real-time data. Therein, city residents and visitors are claimed to live more easily (i.e., more conveniently) due to the integration of these "smart" technologies (i.e., sensors, computing, and automation systems). In terms of city management and service coordination, "smartness" also refers to applying information technologies to different stages of planning, designing, building, and operating cities. (Ronkko, 2018) There can be found more than 36 distinct definitions of the "smart city" concept in a current scientific literature. (Stepanek, 2019) A generalized definition of a "smart city" is a city that uses "electronic Internet of Things (IoT) sensors to collect data and then use insights gained from that data to manage assets, resources and services efficiently". (*Smart city*, 2020) The first attempts to define the concept were focused on the connectivity and features provided by information technology for managing various city functions (i.e., electronic government; e.g., smart meters). More recently, the usage of the term has widened in scope to include the outcome of a "smart city", such as sustainability, quality of life, and services to the citizens. (Ramaprasad, 2017) Of course, these outcomes are generally defined within the context of the market-State.

> **INSIGHT:** *A city could be viewed as a living organism that needs to fulfill specific goals for the city to preserve its existence. Note that this is of course a reification of the city, which could lead to misunderstanding. A city is not, in fact, a living thing; it is the humans (and other organisms) inside the city that are living and that which have need. In actuality, the city exists to fulfill a purpose beyond itself, for humanity.*

Ramaprasad et al., (2017) have proposed a Smart City ontology that attempts to connect early 21st century definitions and unify the concept of a "smart city". Ramaprasad et al. (2017) and Stapanek et al. (2019) consider the ontology a better way to organize smart city concepts than a single definition. This ontology defines the "smart city" concept as a function of two main parameters: smart and city. For each parameter, there is a function that explains the dimensions of the parameter. The concept Smart contains Structure, Function, Focus, and Semiotics. The City consists of Stakeholders and Outcomes. Each dimension from "smart" and "city" is sub-defined as a set of components (or classes).

For the "smart" dimension:

1. Structure includes: Architecture, Infrastructure, Systems, Services, Policies, Processes, and Personnel elements.
2. Function includes: Sense, Monitor, Process, Translate and Communicate.
3. Focus includes: Cultural, Economic, Demographic, Environmental, Political, Social, Technological and Infrastructural elements.
4. Semiotics includes: Data, Information, and Knowledge.

For the "city" dimension:

1. Stakeholders is constructed by: Citizens, Professionals, Communities, Institutions, Business, and Governments.
2. Outcomes include: Sustainability, Quality of Life, Equity, Livability, and Resilience.

The following glossary (Ramaprasad et al., 2017) is necessary to understand the whole definition:

1. Smart: Capable of sensing and responding through semiotics.
   A. Structure: The structure required to manage the semiotics.
      1. Architecture: The architecture to manage the semiotics.
      2. Infrastructure: The physical and virtual infrastructure to manage the semiotics.
      3. Systems: The computer, social, and paper based systems to manage semiotics.
      4. Services: The computer, social, and paper based services to manage the semiotics.
      5. Policies: The policies to manage the semiotics.
      6. Processes: The processes to manage the semiotics.
      7. People: The people responsible for managing the semiotics.
   B. Function: The functions required to manage the semiotics
      1. Sense: To sense the semiotic elements.
      2. Monitor: To monitor the semiotic elements.
      3. Process: To process the semiotic elements.
      4. Translate: To translate the semiotics into action/control.
      5. Communicate: To communicate the semiotic elements.
   C. Focus: The focus of intelligent sense and response (i.e., "smartness").
      1. Social: Social dynamics of the city.
      2. Economic: Economic dynamics of the city.
      3. Environmental: Environmental dynamics of the city.
      4. Technological: Technological dynamics of the city.
      5. Infrastructure: Infrastructure dynamics of the city.
   D. Semiotics: The iterative process of generating and applying intelligence.
      1. Data: The symbolic representation of sensations and measurements.
      2. Information: The relationship among data elements.
      3. Knowledge: The meaning of the relationships among the data elements.
2. City: A city capable of intelligent sense and response.
   A. Stakeholders: Those affecting and affected by the city.
      1. Citizens: Citizens of the city.
      2. Professionals: The professionals of the city.
      3. Communities: The communities of the city.
      4. Business: The businesses of the city.
      5. Governments: Federal, State, and Local governments.
   B. Outcomes: The desired outcomes of a city.
      1. Sustainability: Sustainability of the city.
      2. Quality of life: Quality of life of the stakeholders.
      3. Livability: The livability of the city.
      4. Resilience: The ability of the city to recover.

Herein, the "smartness" of the environment affects the "city" in which people live. The expression between classes of the dimensions "smart" and "city" is:

> **Smart** *(Structure [+] Function [+] Focus [+] Semiotics) [by/from/to]* **City** *(Stakeholders [+] Outcomes)*

Therein, a "smart city" is compound function with two parts/dimensions:

> *Smart City = f (Smart + City)*

The "city" is a function of stakeholders and outcomes.

> *City = f (Stakeholders + Outcomes)*

The "smartness" of a city is a function of structure, function, focus (direction), and semiotics (information processes):

> *Smart = f (Structure + Function + Focus + Semiotics)*

Ramaprasad et al., (2017) define 'semiotics' as, the iterative process of generating and applying intelligence. Semiotics forms the core of the "smart" dimension, such that all other classes of this dimension refer to it. The direction of "smartness" is the "outcomes", which are of interest to the "stakeholders". The direction of "smartness" depends on the "structure" and "functions" of the systems for semiotics. The iterative "semiotics" process, involves data that are converted into information, information to knowledge, and the knowledge is then translated into smart actions. The "focus" of "semiotics" are the relevant possible sub-conceptions of the society and city. The semiotics of each focus will affect the corresponding smartness of the city, its stakeholders, and the corresponding outcomes. The "structure" and "functions" of a city's semiotics (i.e., data, information, knowledge) information system (or, management/coordination system) will determine its "smartness". Together, the four left dimension of "smart" are concatenated to form the "smartness" of a city. Taken together, there are 7*5*8*3*6*5 = 25,200 potential components of a Smart City encapsulated in the definition/ontology. A truly "smart city" is one that has realized a significant portion of these potential components.

Four concatenations are listed below as an example of the 25,200 possibilities:

- Architecture to sense economic information by/ from citizens for quality-of-life.
- Systems to process environmental data for livability.
- Policies to communicate technological knowledge [by professionals] for resilience.
- Processes to translate political information to citizens for sustainability.

> **NOTE:** *This ontology and the following functions are bounded by the conditions and conceptions of a market-State society, and are not entirely representative of a "smart city" in a community-type society.*

Stapanek et al., (2019) note three additional papers that facilitate further explanation of the "smart city" ontology as a function. Babar and Arif (2017) propose a functional ontology with several layers. The first layer is an architecture for planning and decisioning. The second is data acquisition and aggregation, mainly using IoT components, and the last uses pre-processed data for taking decisions and communicating events to citizens:

f (f (architecture
+ (monitor, process, translate, communicate)
+ urban + data)
+ f (citizens + quality-of-life))

Uribe-Perez and Pous (2017) propose a communication architecture inspired by a human nervous system. The architecture is composed of:

1. A sensing layer containing a sensor network.
2. An access layer with "smart" gateways to process a low-level information and act consequently.
3. A data layer with sufficient (e.g., 3) types of databases to store data.
4. A platform layer to supervise and manage the city.
5. An application layer to provide services.

The ontological function is:

(f (architecture
+ (sense, monitor, translate, communicate)
+ urban
+ (data, in formation))
+ f (stakeholders + resilience))

Chen et al., (2016) propose an automotive sensing platform used in the city to obtain data from different parts of the city by cars equipped with sensors. The ontological function is:

f (f (platform
+ (monitor, process, communicate)
+ data))

1.1.3.1 A smart city ontology under community conditions

A glossary for a similar ontology to Ramaprasad et al., (2017), but applicable for community conditions would resemble (differences are underlined):

1. Smart: Capable of sensing and responding through semiotics.
   A. Structure: The structure required to <u>coordinate</u> the semiotics.
      1. <u>Information technology: The software and hardware to coordinate</u> the semiotics.
      2. <u>Projects</u> (services): The <u>projects</u> (services) to <u>coordinate</u> the semiotics.
      3. <u>Teams</u>: The <u>teams</u> responsible for <u>coordinating</u> the semiotics.
      4. Processes: The processes to <u>coordinate</u> the semiotics.
      5. <u>Procedures</u>: The <u>procedures</u> to <u>coordinate</u> the semiotics.
   B. Function: The functions required to coordinate the semiotics
      1. Sense: To sense the semiotic elements.
      2. Monitor: To monitor the semiotic elements.
      3. Process: To process the semiotic elements.
      4. Translate: To translate the semiotics into action/control.
      5. Communicate: To communicate the semiotic elements.
   C. Focus: The focus of intelligent sense and response (i.e., "smartness").
      1. <u>Resources</u>: Resource dynamics of the city.
      2. <u>Access</u>: Access dynamics of the city.
      3. <u>Social</u>: Social dynamics of the society.
      4. <u>Decision</u>: Decision dynamics of the society.
      5. <u>Lifestyle</u>: Lifestyle dynamics of the society.
      6. <u>Life support</u>: Human life dynamics (or, services) of the city.
      7. <u>Technological support</u>: Technological dynamics (or, services) of the city.
      8. <u>Exploration support</u>: Human exploration dynamics (or, services) of the city.
   D. Semiotics: The iterative process of generating and applying intelligence.
      1. Data: The symbolic representation of sensations and measurements.
      2. Information: The relationship among data elements.
      3. Knowledge: The meaning of the relationships

among the data elements.

2. City: A city capable of intelligent sense and response.
   A. Stakeholders: Those affecting and affected by the city.
      1. Users: Users of the city.
      2. Teams: The developers and operators of the city.
   B. Outcomes: The desired outcomes of a city.
      1. Values: Values of the society.
      2. Fulfillment: Fulfillment of the stakeholders.
      3. Flourishing: Flourishing of the stakeholders.
      4. Quality of life (well-being): Quality of life of the stakeholders.
      5. Flow: Flow of the stakeholders.

Four concatenations are listed below as an example of the 6,000 possibilities:

- Information technology to sense life support information by/from users for quality-of-life. Sensors and surveys to sense the quality-of-life of the users of the city and make the data available to users.
- Projects to process resource data for fulfillment. Projects to determine water pollution levels and warn users and teams when they exceed acceptable thresholds.
- Procedures to communicate knowledge for flourishing. Procedures (e.g., notifications) to share knowledge about technological changes to the city with various teams.
- Processes to translate decision information to teams for values encoding. Processes (e.g., optimization algorithms) to translate the social values of the users into decisions that may affect the sustainability of the city.

Using the same expression as Ramaprasad et al., (2017) between classes of the dimensions "smart" and "city", there are 5*5*8*3*2*5 = 6,000 potential components of a "smart city" encapsulated in this definition/ontology. Note that this figure will be off due to the outcome measures not being completely elaborated. For instance, the values are not delineated herein. The market-State is a significantly more complex and convoluted environment than a community-type society, which is why there is such a significant difference between Ramaprasad et al., combinatorial figure of 25,200 components and the 6,750 components of community.

1.1.3.2 Commercial smart city software solutions

There are an increasing number of software and hardware solutions designed to facilitate smart city development and operations, these include but are not limited to:

- IBM: *Smarter City*
  - *Intelligent Operations Centre for Smarter Cities*
- Urbotica: *City Operating System*
- Microsoft: *CityNext*
- Rio de Janeiro: *Centro de Operações*
- Barcelona: *City OS*

2 City design in community

INSIGHT: *A community-city is a whole-expression of our humanity.*

Generally speaking, there are many factors to account for in city design. Sky, atmosphere, and orography generate variable meteorological conditions over a city (e.g., could covering shade, solar, wind, rain, etc.). Buildings and other standing objects/structures induce field dynamics in the environment, including but not limited to: masks, reflections, absorptions, re-emission, rainwater runoff, etc. Human activities produce issues, including but not limited to: heat, noise, waste, pollution, etc. And, human needs (and wants) produce requirements, including but not limited to: life support, technology support, and exploratory support.

Community cities employ the scientific method, prioritizes efficiency throughout appropriate design, have a cooperative versus competitive social structure, are high tech and highly automated, and are the result of a systems approach in managing its complexity. Such cities are a world benefiting platform for the sustainable advancement of humankind. Community cities are circular (generally), fully sustainable, appropriately functional, and access-oriented environments, built for those who are actively engaged in living their life to the fullest.

Cities in community are entirely open source. The result of this openly sourced way of living is that there is the maximization of everyone's potential quality-of-life, and neither hoarding nor fighting over ownership. Community-based cities are operationalized to be continuously up-to-date with humanity's knowledge about how to live more optimally, while drawing upon humanity's inherent and individual strengths. Therein, individuals experience a space where knowledge is applied for the well-being and benefit of all. A lot of the work in these cities has been automated to free up time for individuals to pursue their passions and greater interests. Herein, automation and technology is intelligently integrated into an overall holistic socio-economic design, which primarily functions to optimize the quality-of-life of every individual.

It is possible to have a network of sustainable city systems where humanity has intelligently organized free access to that which is needed so that everyone may thrive; in contrast to an unstable living arrangement where individuals exchange artificial intangibles that everyone is coerced into acquiring and using for [at least] their mere survival, generating socio-economic inequality and the vast number of public health issues that are causal consequences therefrom.

Cities in community provide free access to all goods as services, as in nature, so that individuals don't become constrained (limited) by the abstract intangible known as "money", and hence, disconnected in their ability to accurately sense and appropriately respond to environmental signals. If people have access to the necessities of life they don't "steal", and "crime" (as it is known in early 21st century society) is rendered almost non-existent. The notion that things are "free" in community is something of a misnomer, because there is no money in community. In fact, community can only emerges in a world where everything has been coordinated to be accessible without the need for exchange (i.e., without 'trade').

It is possible to design cities where it is more enjoyable to walk or bike, than to drive, thanks to the intelligent and integrated layout of the physical environment. Among community, individuals walk through the majority of their beautifully planned daily life-space, wherein, they experience a living socio-economic system structured to coordinate decisions, and the flow of resources, for their fulfillment. Therein, individuals experience intentional design that supports a high quality-of-life for themselves and all others; it's an environment where technology and economy serve humanity and the ecology, not the other way around. It is an environment where human creations provide everyone with an abundance of access to life enriching opportunities, maintaining a support structure for living better lives - lives in alignment with the development of individuals' true potential. It is an environment that draws out the best in each individual; pulling out from them the energy of happiness, well-being, and deeply felt love and connection for one another and the universe. These cities are designed provides vast opportunities for outward exploration, as well as the space for humanity to go inward and experience states of universal being. Here, decisions and actions entangle one another in a direction commensurate to humanity's highest potential.

Community-based cities are designed so that there exists the values of efficiency and effectiveness in the fulfillment of human needs, wants, and preferences. Food, energy, transport, and production, for example, have efficiency as a core priority in their designs, which is a necessity for the sustainability of complex socio-technical systems. Material and service constructions are designed to meet human requirements in the best possible manner with the least usage of resources and effort. Conversely, in a monetary system, such designs are generally too expensive. The costs of trying to create a sustainable and efficient city inside a for-profit paradigm are simply too high, which is one of the many reasons there is not a single city optimized for human well-being in early 21st century society. There is very little that is sustainable in how cities in early 21st century society are designed, or the monetary and authority driven social values that have been adopted by their constituents.

Community-based cities are, in general, designed such that food cultivation and natural beauty are integrated into all appropriate and desirable spaces. As a city, community is a place in which all of the tasks (i.e., "jobs") are actually worth doing. Because of the encoding of the value of transparency, everyone in the city knows what needs to be done, and can contribute to the system's continuation and evolution in a coordinated and planned manner.

In community cities, individual's time is their own ad is not structured by coercive structures and authority figures. Here, opportunities for access, self-growth, and contribution are ever present. Individuals' contributions directly benefit everyone, as opposed to working for the direct benefit of someone else or some specifically competitive organization. All work (as effort applied toward the community's continuation and evolution) is relevant, and everyone "owns" their own time.

Due to the intelligently planned design of these cities there are no "prime locations" (as there are in market-State cities); instead, everyone has access to a prime location. Often, it is possible to walk around the living environment and freely pick a variety of flavorful and nutritionally dense foods without worrying about pollution and other toxic residues. These cities feel

> **QUESTION:** *How would it feel to live in a place constructed to express conditions of interest in your well-being as well as facilitate empathic concern for the well-being of others? It may feel like a city that has been designed openly, by all of us, and for all of our well-being.*

Neither the market nor the State has been encoded in a community-type city, and therefore, there is no revenue, no taxation, and no materialization of an environment focused around competition, ownership, and authority. Living in the early 21st century involves (and, for most people it requires) property ownership, and there are taxes and other fees that go along with that ownership. In order to have access, that sort of socio-economic arrangement necessitates either having a job to pay for things, or becoming a ward of someone else who pays for those things. Of course, cities in early 21st century society consequently look and feel very different than they do in community. In the market-State, cities are products and the people within them have little choice but to work for a boss, go on the dole, or starve. Oddly, there is a segment of this population that believes they have something they call "freedom of choice". What they actually have is the illusion of choice, because the options from which they can "choose" have already been decided upon by the structure of the system itself and the "decision makers" higher up in the socio-economic hierarchy or in the distant past; and, these pre-selected options are inescapable if survival is desired.

In community, there is no commerce, no economic trade or exchange of goods, no socio-economic classes or hierarchy, no politics, no bureaucracy, no police, no prisons, no trash, no poverty, no homelessness, and no congestion. When arriving in community from early 21st century society, there is a sense of relief that these things that have held humanity's potential down for so long are no longer present. And still, community creates a city where children and adults alike play outside safely at any hour. Therefore, community is an environment that notably lacks any and all advertising and marketing, in both the physical and digital space. There is no surveillance or misinformation, which are present almost everywhere in cities in the market-State. And yet, the city looks beautifully up kept, it is intelligently laid out, and as individuals move about they don't have to worry about walking on grass or other surfaces that have been sprayed with various killing substances, such as pesticides and herbicides. In community-based cities, no one has to wash industrial pollutants off of their food, or personally filter their water to remove pharmaceuticals, commercial by-products such as sodium fluoride, and other industrial contaminants.

> **INSIGHT:** *Among community, we have a saying, "Systems are what they produce, not what we wish them to produce."*

Individuals in early 21st century society have become habituated to the constant stimulus of commerce and advertising, which wears down (i.e., wears away) their sensitivities to their own needs and their environment. Cities in community are notably void of trash and other pollution. Over time, such pollution causes individuals to turn off from environmental stimuli. The continuously hostile environment of market-State cities causes people to not want to feel their sensory inputs. And, that is the weirdest thing to imagine, that you have to stop perceiving your environment to keep yourself sane. Of course, the light pollution in early 21st century society affects people's sleep, their circadian rhythms, and it prevents them from seeing the stars, which would otherwise provide them with a nightly connection to the larger universe.

In community, the living environment itself almost feels like a single self-regulating and self-healing organism. Community is similar (in this respect) to the human body, which wants to feel well and heal, but needs the correct inputs as well as minimal interference from that which is malignant. Community is a type of society run so efficiently and with organized care that it feels like it takes care of itself. All of those things that are essential for individuals to survive and thrive are integrated and engineered into a unified habitat service system, a city. A city that mirrors the operation of our natural world, which is itself a collection of integrated systems.

> **QUESTION:** *What would society look like if it inherited those properties of the universe that we see as it's incredible harmony and mathematics and self-organization? And, what would it look like if our intention for its creation was to be of benefit to the individual, of benefit to the social, and of benefit to the planet (and even, possibly, the very universe itself)?*

2.1 The integrated city system

> *A.k.a., Total city system, intelligent city system, unified city system, cooperative city system.*

All city systems in community are integrated city systems.

An integrated city system is, in part, in one in which the informational and physical systems of the city are accounted for together in the design and operation of the city. Additionally, integrated city systems account for the total system state of city in relation to its inhabitants and their requirements of the city. Information unification (i.e., looking at society and its cities as an information system first) allows for true cooperation among individual humans and the organization of an integrated living environment.

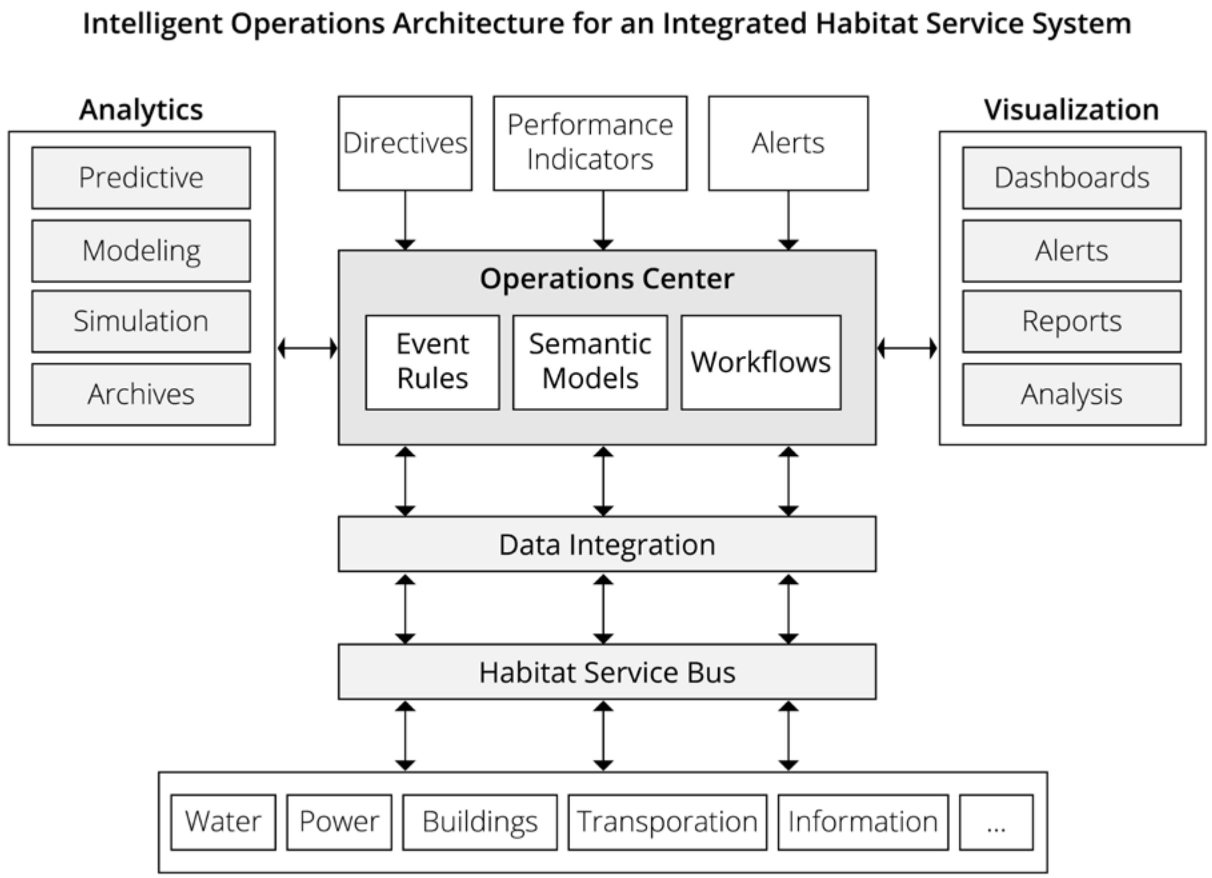
In order to create a life radius that fulfills humanity's real world needs and requirements, cities in community are designed in an integrated manner, and hence, they are often referred to as "integrated city systems" or "total city systems". An integrated city system (a.k.a. total city system) is a city in which every element operates together efficiently as a whole system. In other words, all aspects of the construction and functioning of a community-type city are well integrated. Instead of leaving city functions under the control of isolated organizations, individuals and obscured programs, cities in community integrate their control. All functional aspects of these cities, from food cultivation to sewage and energy production are processed together as one system (i.e., they are 'integrated'). Integrated city system are strategically planned and cooperatively operated socio-technical environments. In order to accomplish their functions, integrated city systems are data collection and product materialization platforms.

City systems in community are integrated living environments. Significantly, said city systems are [sustainably] integrated into the planetary biosphere/ ecology by means of a planned and bounded service area known as a habitat service system. Therein, these cities are integrated into the form of a human habitat through intelligent socio-decisioning and thoughtful material and lifestyle design. Community cities are based on individuals who play an active part in their design and operation.

An integrated city system is one in which:

1. The elements of the city are interconnected (informationally and spatially).
2. Access to all parts of the city exists through continuity of travel by various modes of transport.
3. The city layout exists by reason of function, demand, and optimization (i.e., function, effectiveness of fulfillment, and efficiency of fulfillment).

Figure 5. *Depiction of a city-level operating system for InterSystem Team Operations.*



In community, ideas and designs are well thought-out and coherently integrated into a unified information system [model] before being encoded into decisioning and constructed into the environment, whereupon they are tested to ensure desired alignment. A total city system approach requires systematic design and overall planning to attain a high standard of living for all the occupants.

> **NOTE:** *A total city system is a city that accounts for the total (whole) environment.*

It important to address an issue here: the notion that intelligent core-systems planning, implies mass uniformity, is not accurate. Cities in community would be uniform only to the degree that they would require far less materials, save time and energy, and be flexible enough to allow for innovative changes (through modularity), while preserving the local ecology. Cities in community are planned so that they are capable of fulfilling the needs, wants and preferences of all community inhabitants. Through planning and testing we are able to produce a pleasant and desirable living space that removes urban sprawl and can effectively account for social, economic, and ecological problems. The integration of function is necessary for the optimization of our fulfillment, as well as an accountable solution-orientation to any problems that may arise.

Herein, information processing and automation systems are combined with sensors and human effort (where necessary and/or desired) to optimize the operating efficiency of the city. The use of up-to-date technological methods, including electronic feedback, digital information processing, and automation, is applied to the entire city system. The use of automation ensures that what we intend to happen, actually does happen, every time we want it to happen. Through the application of computing we are able to process trillions of bits of information per second, which is useful (though not absolutely essential) for the facilitation of complex multi-variate decisioning, and hence, the coordinated operation of these cities. Intelligent coordination keeps a city's services operating at peak efficiency and uptime, maintaining our materially desired fulfillment, and creating an optimized economy that avoids overruns and shortages. For example, the irrigation and fertilization of a primary food cultivation belt (within one of these cities) is programmatically controlled through an automated irrigation system involving environmental sensors, integrated circuitry, and various mechanical technologies. Hence, the emergence of a service system that frees humans from unnecessary labor, makes the most efficient use of resources (water in particular), while ensuring a sustained healthy landscape. Waste management, energy generation, and other services are managed by these "smart" (i.e., "cybernetic") methods. This integrated control is openly programmed by us, for us (as a community), and applied throughout these city systems for social and ecological concern.

Additionally, an integrated city system is also defined by the consolidation of as many functions as possible (or desired) into the least amount of material area. For example, most of the outer surfaces of buildings convert solar energy into electricity, and the surfaces are themselves fitted with automated cleansing systems.

Integration, not only within a city itself, but between cities provides innumerable benefits, including but not limited to:

1. Increased ability to identify problems.
2. Increased ability to aggregate information and identify useful/applicable information.
3. More informed responses.
4. One platform provides better coordination.
5. Better communication and cooperation.
6. Increased safety.

The total societal system may be delineated as follows:

1. One solar and planetary system.
2. One unified societal system design [specification standard].
3. Four societal information sub-systems (social, decision, material, lifestyle).
4. One global habitat service system (network of city systems, the economic global access system).
5. The local habitat service systems (individual, integrated city systems).

2.1.1 The cybernetic city

> *A.k.a., The cybernated city, the computationally integrated city, the diagrammatic city, the smart city, the computational city, the intelligent city, the automated city, the computed city, the city operating system, the urban operating system, the city information system, the city central processing system, cyber-physical-social systems (CPSS).*

Cybernetics is an interdisciplinary science for exploring digital, mechanical, or biological regulatory systems. (Wiener, 1948) Classical cybernetics has evolved, since its instantiation in the 1940s into second (or, third) order cybernetics-the cybernetics of observing (includes the observer, rather than only being observed) systems, which also concerns the principles of learning and communication. (Ross, 1957; Glanville, 2007) In line with the original meaning of the word, "to steer, navigate, or control", cybernetics can be applied to the design and operation of any complex system, including cities. (Ronnko, 2018)

> **NOTE:** *The central element(s) in cybernetics is control, which implies feedback, decisioning, and communication, which all involve the transmission of information.*

At the core of a cybernetic city (a.k.a., smart city) is a software control (i.e., cybernetic) center in which planning and operations occur. Since the science of cybernetics focuses on the information and communication involved in the process of feedback and decisioning, a cybernetic city is one in which the latest information and technology is utilized to maximize the fulfillment of the inhabitants. (Lasker, 1981) At the core of cybernetic city operations are clear, real-time situational information for monitoring, analyzing, understanding, planning, and operating smart cities.

> **NOTE:** *The idea of a cybernetic city and the emerging conception/ontology of a smart city are highly related, except that the cybernetic view is more systems- and community-based, whereas the notion of a smart city is more based in the market-State ideology.*

A cybernetic city is a cyber-physical-social system. Common cyber-related terms to understand this complex relationship include, but are not limited to:

- **Cyber-social systems (CSS)** - social systems with embedded digital structures and devices to facilitate human scale endeavours. Cyber-social systems are a collection of technologies for coordinating and controlling interconnected social and computational capabilities. CSS is the merger of cyber (electric/electronic) systems with social structures.
- **Cyber-physical system (CPS)** - is a system featuring a combination of computational and physical elements, all of which are capable of interacting, reflecting and influencing each other. Cyber-physical systems are a collection of technologies for coordinating and controlling interconnected physical and computational capabilities. CPS is the merger of cyber (electric/electronic) systems with physical things. (Trappey et al., 2016) For example, mechatronic systems, which combine the disciplines of mechanical, control and electrical engineering. CPS systems include automated systems that sense and control physical phenomena through sensors, processors, and actuators.
    - **Human-in-the-loop (HiTL) CPS** - CPS systems that involve control loops with human goal-oriented interaction.
- **Cyber-physical-social systems (CPSS)** - the integration of cyber space, physical space, and social space.

Fundamentally, cybernetics is the science of self-regulating systems, which (1) exist in living matter and its relationship to its environment, (2) as the interaction among living things, (3) in machines, and (4) in the interaction between living things and machines. In the context of cybernetics, self-regulation includes processes that maintain organisms or organizations as viable entities and that enable machines to perform selection and control operations. (Lasker, 1981) A city system that is systematically involved in the production, organization, distribution, and use of knowledge and information, which constitutes a self-regulating city system.

Amstutz (1968: 21) states that a city could be made more response to its populations needs via a threefold strategy:

1. Structuring the environment into categories and subcategories. For example, identifying the core services of the habitat (e.g., life, technology, and exploratory).
2. Developing clear objectives and criteria for evaluation. For example, what is flourishing, fulfillment, and quality of life, and how are they measured.
3. Using computers to 'synthesize and maintain a representation of the total environment'.

Amstutz's approach rests on the delegation of control ("authority") to computer systems. If city functions were pre-programmed, then city planners and operators would be able to approach city problems with "'increased effectiveness due to the availability of more meaningful data and an increased (model based) understanding of [the] environment" (Amstutz, 1968: 21).

> **INSIGHT:** *Intelligent systems evolve through feedback phenomena. Feedback is an essential action in the generation of a sustainable city environment, for both the efficient use of resources and the integration of effective functionality.*

Computer science, systems science, and simulation are early sources of inspiration for viewing the city as an operating platform for humanity. Since the 1950s and Norbert Wiener's laying out the principles of cybernetics, the city has increasingly come to be viewed as a communications system. (Meier, 1962; Webber, 1964; see also Light, 2003) The city is a space of data flows and environmental modelling is traceable to the digital computation work of Forrester (1961; 1969). Forrester thought of cities from a scientific perspective (as in, the 'science of cities'; Batty, 2013; see Townsend, 2015 for a critique), and saw the city as a complex (yet arguably linear) system of interacting parts experiencing growth, equilibrium and stagnation; a system easily modelled through calculated flows and an account of conditions in the surrounding environment. Batty et al., (2011) state that, "One of the key differences between theories of cities developed a half century or more ago and the emerging science of cities and societies in the early 21st century revolves around the idea that the focus should no longer be on location, but on interactions and connections, on networks and the concomitant processes that define flows between places and spaces." The understanding that computer applications, system dynamics and digital modelling are mechanisms to solve societal, and particularly, city problems was espoused by a generation of planners and technologists, one of the most notable being Jacque Fresco who envisioned architectural structures (and even, whole city systems) optimized for and by computer aided environments. (Fresco, 2007)

An cybernetic city system establishes a diagrammatic form of relationship with the city. Diagrammatic modeling and simulation occurs for information structures and dynamics as well as spatial structures and dynamics. Informational diagrammatic control involves the visualization of information to easily identify functions and simplify decision selection. In effect, a unified city operating system establishes a diagrammatic form of relationship with the city. (Marvin, 2017: 92) Fundamentally, cities can be visualized, diagrammed, and all aspects can be simulated. Therein, coordination and control of the [cybernetic] city is given over to computational logic-- involving the coordination of information and material flows through information systems and technologies, and their interface with the material world.

INSIGHT: *The conception of a cybernetic city carries with it the possibility of information-system-based planning and cybernetic coordination. Cities can be known, planned and controlled in large part through data processes and algorithms.*

In a cybernetic city functions that are kept separate and loosely coupled (e.g. waste collection, transport provision, energy services, security and emergency response) are planned and operated in an integrated relationship. Therein, there is a single, unified information system that accounts for software and hardware systems that interoperate and are interconnected.

NOTE: *In the industrial environment, enterprise resource planning (ERP) systems have been used extensively to coordinate the flow of resources in order to streamline internal operations, linking finance, procurement, payroll and human resources in cities. ERP systems effectively render internal resources relations predictable and controllable. The use of ERP implies a functional understanding of the organization, wherein the division of operations into functions and sub-functions is crucial for the appropriate functioning of the whole. In an ERP system, organizational operations are detailed as a breakdown of components into sites (locations), agents (subjects), functions, and relationships. Note that there are also resource planning systems for various sub-industrial functions, including, for example, manufacturing resource planning (MRP).*

Technology embodies routines and procedures that generate particular forms of perception and cognition, both shaping behaviour due to the processes of functional simplification and reification by which a prescriptive order is formed (Kallinikos, 2011:7). Kallinikos explores different techniques of coding with a particular focus on object-oriented programming. Object-oriented programming is a structured form of software coding, organized by structures and procedures, that divides reality into objects, which are further divided into sub-objects. Each object has attributes, and by recombining attributes the relationships between objects can be reconfigured. This computational logic renders reality as a set of integrated information, and thus, usable in the real world.

Through an emphasis on modularity, along with pre-determined structural features and intrinsic qualities, information technology packages and knowledge are constituted as both specialized and transferable--from or organization to organization or city to city (Voutsina et al., 2007).

A cybernetic city, as a system of systems, operates through techniques of classification, resulting in the provision of a system for organization and, in this way, a framing for an objective reality. This classification process involves the development of typologies, the establishment of system hierarchies and a mapping of connections between these components. Such direct identifying and explication of interconnectedness renders the entire system of internal relations predictable and controllable. Classification also has an ontological function, by determining components and establishing a set of relationships, thus creating entities with definable boundaries (e.g., service inputs, processes, and outputs). The integrated visualization of a city as a system of systems necessarily involves the development of a detailed map for organizational action and control. (Marvin, 2017: 93-94)

Current hardware and software technologies allow for city-scale operating systems. The idea of the city as an operating system has been discussed in the literature. First, it has been used as a 'metaphor' in which cities are seen as interchangeable with computer systems. (Marvin, 2017) Therein, the city is viewed as an information [processing] system based on the acquisition, storage, processing, and retrieval of information and materiality – an operating system. Through these information technology systems, locations and actions are capable of being sensed in real-time, wherein the operating system aggregates and processes data leading to decisions and actions at a distance. (Marvin, 2017). The resulting 'real-time city' operates through sensor networks, computing frameworks, and automated hardware that aggregate data streams into services and products (i.e., fulfillment) for their users (Townsend, 2000; 2015). Further, Easterling (2014: 5) examines how a combination of infrastructure space, sensors and software are may be designed to use information to "determine how objects and content are organized and circulated [in] an operating system for shaping the city". Easterling (2014:6) describes an operating system as a platform, both updated over time and unfolding in time to handle new circumstances and situations, which uses software 'protocols, routines, schedules and choices' to encode relationships between buildings or managing logistics of infrastructures. This later view describes the operating system as a platform for city control. Thus, it is possible to view, understand, and operate a city through an examination of the hardware and software systems that coordinate and control its behavior.

The software components of these systems include, but are not limited to:

- Databases, predictive systems, analytics, modelling and simulation.

The hardware components of these systems include, but are not limited to:

- Computers, sensors, control rooms) assembled into purpose-built platforms for functional and spatial integration.

NOTE: *In practice, these software and hardware systems form a hybrid of techniques, tools, and software systems.*

Responsive city design and operation is not just about the convergence of different technologies, it is also about the convergence of semantic structures (perceived environment and life world) and syntactic structures (services and infrastructures) over time. Thus, cybernetic cities can be viewed from four primary dimensions:

1. **The conceptual cities (1D-cities)** - for example, the conception of a city
2. **Bi-dimensional cities (2D-cities)** - for example, GIS overlay data.
3. **Three-dimensional cities (3D-Cities)** - for example, 3D mesh models of a city's objects.
4. **Dynamic, spatio-temporal cities (4D-Cities)** - for example, a simulation of the city with mesh models and GIS data over time.

Cybernetic modules for a city system are likely to include at least three principle elements (Costa, 2019):

1. **Instrumentation** - the ability of systems to measure information by means of sensing tools. Instrumentation is the first movement of action against entropy. Examples are locative media (LBM), georeferencing (GIS) and remote sensing. Herein, environmental, energy, and social sensors serve as parameters.
2. **Analytics** - informatics to perceive and interpret acquired information in accordance with a set logic. When a system acts upon information within the information environment to produce more useful information. In city design, parametric methods include: BIM (Building Information Modeling), SIM (System Information Modelling) methodologies, and performance management (PM).
3. **Actuators** - when a systems acts physically within the spatial environment.

NOTE: *Instrumentation and control systems are used to automate processes. Items to be included in the design and analysis of these systems are: reliability of control of critical processes, safety of personnel, and suitability of instruments and control devices in the environment in which they are installed.*

Cybernated habitat data collection sources (terminology) and assessments include, but are not limited to.

1. **Environmental monitoring networks** - Networks that provide data [sources] on environmental variables. For example, weather data, air quality, user health data, etc.
2. **Fixed and mobile sensory arrays** - arrays of sensors, such as those attached to the interior or exterior of buildings, or airbone platforms.
3. **Real-time sensors** - sensors that collect and transmit data to be processed in real-time.
4. **Recorded sensors** - sensors that collect and store data to be processed at a later time.
5. **Distributed sensory networks** - sensors to monitor and collect data on physical phenomena, physical conditions, and physical systems.
6. **Biomonitoring (biological monitoring)** - the assessment of an ecosystem base on organisms living in it. The lives in the ecosystem.
7. **Data contributed by city inhabitants (crowd sensed, social)** - city inhabitants articulate issues and other data.
8. **Global positioning system (GPS) -** data from a satellite-based global positioning system.
9. **User profiles** - the current, check-in, or modification of user profiles.
10. **Habitat assessments** - the assessment of an ecosystem based on its physical characteristics. The physical characteristics of an ecosystem.

The four characteristics of city-level data for a cybernated system (i.e., "big data") are (Santana, 2017):

1. **Volume:** coming from many data sources distributed across the city.
2. **Variety:** data is collected from different sources, and have structured, semi-structured, or unstructured formats, such as video records, relational databases, and raw texts, respectively.

This is important for cities, because city data is collected from multiple sources.

3. **Velocity:** data processing must be fast and, in some cases, real-time, or it may be useless.
4. **Veracity:** because of the large amount of data collected, and the use of multiple data sources, it is important to ensure data quality, because errors in the data or the usage of unreliable sources can compromise its analysis. In cities, incorrect GPS readings, malfunctioning sensors, and malicious users can be sources of poor data.

2.2 Cybernetic city automated operations control system components

NOTE: *City planning may be otherwise viewed as the pre-programming of habitat [city] functions.*

A cybernetic operational control system for a highly automated city would necessary involve:

1. A project coordination system, including but not limited to:
   A. A tasking system with tasking flow automation
   B. A documentation system with document flow automation
2. A unified information database
3. A unified information coordination system (a.k.a., information management system, IM; e.g., BIM)
4. Continuous system design and development software
5. Models development to control devices and facilities (e.g., buildings)
6. Software and hardware (hybrid) systems to organize, coordinate, and control operations.

The conception of an operational information model of a cybernetic city control system requires solutions to the following tasks. In other words, the following tasks must be solved for the functional conception of an operational information model of a cybernetic city (Kuzina, 2019):

1. **Task 1:** The identification of alternative technical systems that implement the goals and objectives (i.e., a probabilistic decision system).
   A. **Input information:** Tasks and criteria, general requirements for the technical complex (product), the composition of the complex and the requirements for subsystems, the approximate terms of use, the data of scientific and technical information.
   B. **Output:** Principles of design solutions, the required technology and materials, the required solutions and scientific and technical problems, the tree of alternative versions of the technical complex with an assessment of the existing state of availability for each of the options and an assessment of the probability of creating a technical complex to given estimated time
2. **Task 2.** Full assessment of alternative solutions and selection of the solution according to the objectives criterion.
   A. **Input information:**
      1. Product characteristics (for each alternative): static characteristics (e.g., product design specifications, weight, geometric dimensions), dynamic characteristics of the product in different operating modes.
      2. Characteristics of the product life cycle: the required volume of production works for the solution; a calendar date of completion of research, development, production and operation; the duration of the stages of development and production (standards of the times); the economic parameters for life-cycle stages of complex (cost standards).
      3. Criteria and models of the target effect/ outcome.
      4. Model and objectives criteria.
   B. **Output information:** Evaluation of the objectives criterion for each alternative solution; comparative characteristics of alternative solutions for different parameters; reasoning/ justification of the proposed solution.

NOTE: *For this purpose it is necessary to create specialized software that can provide decision support for each task level.*

The composition of a system-wide mathematical and software automated control system operation of the object is divided into 4 subsystems (Kuzina, 2019):

1. **Message/signals analysis system** - determination of incoming information processing modes and providing necessary dialogue between users and technical means. The mode of such messages processing should be determined by the system on particular features of messages. There are five modules (or, blocks) for this system:
   A. **Module coordinator (dispatcher module; 1)** - designed to ensure the joint operation (co-operation) of all units of the message analysis system in accordance with the type of message, the configuration of the system in accordance with the allocated resources, and the implementation of communication with other systems involving mathematical support.
   B. **Modules of (2) syntactic and (3) semantic task analysis** — the allocation of individual

sentences of messages, checking the correctness of their construction, their distribution, in order of importance, the formation of the summary rules of their analysis, the definition of input and output parameters of the message, the formation of signals in the block dialogue about anomalies identified during the syntactic and semantic analysis.

C. **Module of the works list (4)** - intended for determination of the message processing possibility (transition - from input parameters of the message to output), determination of optimum ways of processing, creation of the list and sequence of works with their necessary description and formation of output arrays structure with their description, formation of signals in the module of dialogue about reception of the message in processing or about impossibility of its processing.

D. **Module of dialogue with the user (5)** - provides formation and delivery to the user of signals about acceptance of the message in processing or impossibility of its processing, about the anomalies revealed during the syntactic and semantic analyses, the analysis of additional (secondary) messages of the user, addressing them in other modules of system and formation of the corresponding signals to the user about implementation of its additional messages.

2. **Information support system for task solving** - The system is designed to organize the storage of information and provide the necessary information to solve all calculation and information problems. The information support system includes 7 modules. Additional modules can be included in the system, such as standard procedures, placement optimization modules, information security, and statistics collection.
   A. **Module coordinator (1)** - to ensure the joint operation of all the modules of the system.
   A. **Module of information requests analysis (2)** - performs the functions of perception, semantic analysis of the request, determining the optimal way of its processing.
   B. **Modules of (3) formation, (4) updating and (5) maintenance of information arrays** (fields) - ensure the compilation of information record structures, the establishment of semantic (associative) links, the compilation of addresses, the location of records, the organization of new data or changes to existing records, the elimination of obsolete records.
   C. **Module of information retrieval (6)** - provides the determination of the location addresses necessary for solving a specific task of information, the selection of information from the corresponding information files, the organization of the primary grouping of information in accordance with the requirements of a specific information request.
   D. **Module of response arrays formation (7)** - determine the form of the response array, which is necessary for solving a specific task, selecting and arranging information in the necessary order, selected and grouped by the search unit, including standard library procedures into operation, which are not explicitly in the main information arrays.
3. **Organization system for task solving** - direct control of a projects (or, programs) set of work including mathematical support at the solution of information and settlement tasks. To perform its necessary functions (*see below*), there are 5 modules:
   A. **Module coordinator (1)** - to ensure the joint operation of all the modules of the system for solving problems, setting the system to work in the mode corresponding to the allocated resources, and communication with other parts of the system involving mathematical support.
   B. **Module for planning (2)** - provides the definition of the resources required to solve the problem and the formation of the corresponding application, the planning work on the solution of tasks when selected.
   C. **Module for task library maintenance (3)** - maintains a library catalogue searches and a call to the required programs, should maintain and update the library and directory.
   D. **Module for control (4)** - ensures the development of the plan of computational work, timely connection to the necessary programs, the formation of appeals to the exchange unit in the case of joint work of several programs.
   E. **Module for exchange (5)** - organizes the joint work of several programs, processing of additional instructions received in the course of solving problems, monitoring the use of allocated resources and the time of return of free resources.
4. **Automatic project coordination/management system** - designed for registration and accounting of all appeals to the system, differentiation of access to information and tasks. To perform its functions (*see below*) the system has 6 modules:
   A. **Module coordinator (1)** - provides for the

joint functioning of all the modules of the system in all modes: applications for inputting, outputting information and solving problems from individual external subscribers, technical personnel of the facility, other automated objects of the system, other tasks solved in the system, etc.

B. **Module of message registration (2)** - registers messages.
C. **Module of checking request authentication (3)** - authentication of a request based on unique characteristics of calls (names or numbers of subscribers, various digital, light codes, especially voices, etc.) identifies subscribers and checks their right to Enter, output information and solve problems.
D. **Module of newly formed information classification (4)** - classification (establishment) block of the newly formed information column automatically, based on a meaningful analysis, determines the right (security classification) of different subscribers to use information, which is a synthesis of individual messages or the result of solving problems.
E. **Module of accessing information organization (5)** - prohibits or allows access to information and tasks without the permitting commands of the authentication checker unit, and organize access to information and tasks with appropriate permissions.
F. **Module of registration and information delivery (6)** - registers the delivery of information.

In concern to the information support system (*see above*), to reduce the volume of operations for the preparation and input of information into the system, eliminating unnecessary duplication of work and information, reducing the required amount of memory and unification of mathematical support, it is necessary to create a single array (fields) of information to solve all problems of automated control systems. According to the efficiency of use and physical storage of information single arrays (fields) can be divided into levels (Kuzina, 2019):

1. Permanent information.
2. Operational information required to solve a set of tasks of one stage of management.
3. Current, information needed to solve a specific problem.

The objectives of the information support system include (Kuzina, 2019):

1. Reception, placement and storage of information.
2. Search on information fields and selection of information necessary for solving specific tasks.
3. Processing of selected information, editing and formation of response information.
4. Arrays (fields) in the form necessary for solving specific tasks.

The following requirements are necessary in order to solve problems associated with the necessity for operational information (Kuzina, 2019):

1. Ensuring efficient and optimal use of data, information and all types of resources.
2. Automation of production processes, decisioning processes in the event of deviations from the planned indicators (or, pre-planned flows).
3. Complex systems formation for interaction of production and socio-technical processes.
4. Ensuring information interaction between people and between people-and-machines as a means of communication and information transfer.
5. Development of a learning system, system of knowledge accumulation and information coordination within the society.
6. Predictive analysis of scientific and technological development, forecasting of engineering systems.
7. Risk assessment and calculation of the probable consequences of adverse circumstances.

In concern to the organization system for task solving, the system should provide the following functions (Kuzina, 2019):

1. Specific planning of computational work required to solve task problems.
2. Determination of the necessary system resources to solve a problem.
3. Timely inclusion in the work of some programs of special mathematical support.
4. Monitoring the progress of the task and its logging.
5. Processing and maintenance of additional instructions received in the course of solving a problem.
6. Definition of capabilities and management of parallel solution of several tasks.
7. Modify the plan of solving the problem and the redistribution of computational efforts in the case of changing the allocated resources.

The operation of the automatic project coordination system should provide the following functions (Kuzina, 2019):

1. Identification of the subscriber who has applied to the system for input, output of information or solution of this or that task;

2. Check the rights of the subscriber to input, the output of this information, and that the solution of a particular problem;
3. The permit input, output information and the solution of the problem or a signal of disloyalty of circulation;
4. Definition of the classification of the newly generated information (the solution of task or generalization of individual messages) on the right of secrecy;
5. Registration of all requests for input, output information, problem solving with indication of subscribers, time entered or issued information.

The requirements of a software system for the operation of the prior detailed automated control system, which can take an effective final decision, include (Kuzina, 2019):

1. Database containing data of statistics reports, goals and requirements of the users of projects.
2. Software modules for collecting information, importing data into the repository by both automated and manual input that depend on the required information and its source, modules for calculating performance indicators and comparing options.
3. Analytical subsystem of standards, infrastructural elements, suppliers of materials, equipment, etc.
4. A planning subsystem for predicting the results of selected solutions, based on the calculations of local problems, which performs calculations in the form of comparison.
5. A means of visualization of the obtained multi-factor parametric models. Means of display of initial data at the stage of information input, results of changes of the main criteria depending on the chosen decision for each parameter, results in General on object. Generation of reports in various formats.
6. An administrative subsystem is necessary to ensure information security (taking into account the differentiation of access rights to information, the order of use of data libraries), to work with database servers.

In addition, the usage of an information system for coordinating the operation of buildings and their infrastructure will allow (Kuzina, 2019):

1. Improve the efficiency of design, construction, operation on the basis of predicting the behavior of the building system and its infrastructure.
2. To organize rational management of the project implementation by increasing the level of operation planning at the initial stages of design and increase efficiency in the implementation of tasks.
3. Build a predictable financing system for the facility throughout the life cycle of the building, simulate changes in infrastructure projects.
4. Reduce time for preparation and execution of works, labor costs for operations on search and processing of data for decision-making.
5. Provide the proper level of security in operation of life support systems in smart city.

Additional requirements for the implementation of a cybernetic city system include, but are not limited to:

1. A universal coding system (or, universal code) for the unique identification of all recorded knowledge and information. The designation of an information system designed to provide global access to all knowledge and information. All material identified by this code can be located by the use of a multi-category index. These categories include, but may not be limited to: (1) subject terms and phrases, (2) proper names, (3) geographic names and places, areas, or segments, (4) type of material, and (5) level of material. (Lasker, 1981)
2. The software for all habitat service systems.
   A. An information system to account for all planning and operational activities at any given time.
      1. Systems to collect, analyze, model, optimize, and visualize operations of city systems.
      2. The software architecture to monitor, process, translate/control, and communicate city and human data.
      3. Information software to acquire and understand environmental data for livability.
      4. Information software (decisioning) to determine optimal materializations from knowledge for resilience.
      5. Information software processes (communication) to translate information to users for sustainability.
3. The hardware for all habitat service systems.
   A. The hardware architecture to monitor, process, translate/control, and communicate city and human data. For example, sensors to sense economic information from users; sensors to sense environmental and ecological information; hardware Systems to computationally process data.

2.3 The network of cities

A.k.a., The cities network, the city network, the geometric network of cities, the global network of cities, the global city network, the community-city

network, the city-community network, networked cities, the city system, polycentric urban configurations.

A city network is a specific type of spatial structure formed through the combination of city agglomeration and connection within and between cities. (Ni, 2017) A community-type societal system is materially composed of a network of integrated city systems that operate together to create a unified, global habitat service system (i.e., a single, global economic/access system). In community, there exists a global network of cities (i.e., a global city network, or global cities network). In other words, said society materializes as a network of integrated city systems that operate through a unified, global habitat service system consisting of all the cities in the [community] network. Cities in community are set within an enormous, global city network connected by the various flows of information and materials.

NOTE: *Cities in a community-city network are both independent (in that they are self-integrated) and interdependent (in that they share access to resources and services).*

Cities in a community city network are unified under a single societal information system. Therein, the network of city systems is represented by the Global Habitat Service System (a.k.a., a true global access system), followed by the local city systems, represented by the Local Habitat Service Systems. Simply, there is one global conception of a service system for global design and accounting, and then, there are many locally materialized city expressions.

NOTE: *The total material system of community operates as a united network of cities with shared, coordinated access to global resources and services.*

An analytical framework for a city network must account for:

1. Centrality (intra-city development) - the design and development of a single city; .
2. Inter-connection (inter-city development) - development and access between cities.

Note that the market-State defines a network of cities as two or more previously independent cities that work toward jurisdictional and/or economic "cooperation" to achieve faster and more reliable trade, transport and communications infrastructure. The evolution of cities in the market-State is toward a network of jurisdictionally interrelated cities that trade with one another (i.e., "trading cities"). (Batten, 1995)

INSIGHT: *A city network can distribute the load of production.*

Further, in the market-State, State/jurisdictional relations and political situations determine whether there is a weak or strong socio-political connection between cities. Therein, global socio-political situations and changes play a significant part in the remodeling of cities, thereby affecting the jobs, wealth, mobility of occupying inhabitants (e.g., mass migration due to unrest in other geographic locations).

2.3.1 The networked grid of cities

A.k.a., The grid of cities, they city grid network.

Generally, cities in community are laid out in a geometric grid-like manner. When viewed from above, cities in close proximity to one another in a community network of cities are often seen to be laid out in a geometric arrangement, wherein individual cities are located at the vertices ("points") of whatever shape the geometric grid of the arrangement takes. For example, cities could be laid out in a hexagonal-like grid structure with cities at each vertex of the repeating hexagonal shape, wherein the transportation network between cities is placed at the edges of the repeating hexagonal shape. In other words, when zooming out from an integrated city systems, there is a visible return to nature before a network of such cities appears in geometric formation, and possibly, clustered.

NOTE: *Each city in the community network is part of a unified community [habitat service] system, and connected via a mass rapid transportation system.*

Frequently, the total global city network is divided to city clusters, wherein many cities are clustered in geometric proximity to one another. However, this clustering arrangement is highly dependent upon geographic region, with clustering not being possible in some geographic regions.

2.3.2 Fulfillment profiles

Cities in community could be viewed as "fulfillment centers". Therein, if "you" don't like a particular fulfillment center (i.e., a particular city), then there are other cities in the community network that may resonate more greatly with "your" fulfillment profile.

2.4 City surface mediums

On Earth, there are presently two surface mediums for city construction:

1. Land-based cities - cities positioned on land.
2. Ocean-based cities - cities positioned on bodies of water.

A global network of cities in the sea can easily accommodate many millions of people and relieve the land based population pressures. On the ocean, ships could act as integrated manufacturing platforms producing products as they travel.

2.5 *City structuring*

> **NOTE:** *The following descriptions attempt to be societally agnostic.*

There are three primary types of city structure:

1. **The radio centric city** - Radiates outward from a common center. Inner and outer ring roads are linked by radiating roads. A direct line of travel for centrally directed flows. A radio-centric city does not have to be a circular city; a square or other shape can also be laid out as a radio centric city. Moscow, for example, is a radio-centric city, with the center of all rings being Moscow Kremlin and Red Square.
2. **The gridiron city (rectilinear)** - Composed of straight streets crossing at right angles to create many regular city blocks. This form is typical of cities built after the industrial revolution. Requires flow hierarchies. Is potentially monotonous. Flexible grid expansion.
3. **The linear city** - A city expanded along a linear transport system. Very sensitive to blockages. Uni-dimensional linear expansion. Dubai and Navi Mumbai are examples.

The most well-known models of city land growth include:

1. **Concentric model** (concentric zone model, Burgess model) - city grows radially outward from a single point. Ideally, different land uses are distributed via concentric rings around the city center.
2. **Sector model** (sector zone model) - city grows sector by sector.
3. **Multi nuclei model** - city grows from several independent points rather than from a central area. Little to no planning; almost completely ad hoc.

In the market-State, rectilinear cities have several advantages and disadvantages given those conditions (Levinson, 2020):

- Advantages include, but are not limited to:
  - Maximizes the use of space for square/ rectangular buildings.
  - Simplifies real estate by making market-State surveying easier.
  - Is embedded in existing property rights, effectively making the property rights structure [nearly] impossible to change.
- Disadvantages include, but are not limited to:
  - Is among the least efficient way to connect places from a transportation perspective.
  - Reduces opportunities for nature, interesting spaces, architecture, etc.
  - Wastes developable space by overbuilding roads.

The market-State urban population-dimension hierarchy generally scales in the following manner:

1. **Hamlet** - may only include a few dozen people and offer limited services. These are clustered around an urban center and may only consist of basic need services.
2. **Villages** - larger than hamlets and offer more services.
3. **Towns** - more urban with a defined boundary, but smaller than a city in terms of population and area.
4. **Cities** - densely populated areas that may include tens of thousands of people.
5. **Metropolis** - large cities and their suburbs.
6. **Megalopolis** (conurbation) - where several metropolitan areas are linked together to form a huge urban area.

In concern to sustainability, the key design issues in city structuring include, but are not limited to (Marshall, 2005):

1. The need to create layouts that minimise the demand for energy and materials.
2. The need to create layouts that eliminate the automobile, while facilitating walkability and utilizing mass rapid transport.
3. The need to use environmentally friendly modes of transport, especially walking and cycling (which may also bring health benefits).
4. The need to space service locations appropriate to demand.

The divisioning of districts/zones in a city occur based upon (includes, but is not limited to):

1. **Demand** (i.e., # of people; mass and volume of demanded objects)
2. **Service type** (e.g., business, residential, recreational, etc.)
3. **Access type** (e.g., personal, commons, system)
4. **Location and proximity** (of people to goods and services)
5. **Socio-economic class** (market only; e.g., high-income class, middle-income class, lower-income class, poverty class)

It is relevant to note here that many of the ongoing early 21st century theories around city design and development are marred by presumptions based on outdated understandings, poor quality data on humans' real-world requirements, and dis-unified modeling. Central place theory, for example, is a market-State geographical model of the spatial distribution of cities across a landscape that sought to explain the number, size and location of human settlements in a residential system, and was developed in 1933. (Ben-Joseph, 2000) (Pumain, 2004) The theory was created by the German geographer Walter Christaller, who asserted that settlements simply functioned as 'central places' providing services to surrounding areas. (Goodall, 1987) (*Central place theory*, 2019)

The sub-conceptions of central place theory are [out-dated and market-focused]:

- **Central places** - urban centers that provide services to their surrounding rural people (hinterland).
- **Threshold** - the minimum number of people needed to support a particular function's existence in a central place. Including, the minimum number of resources and tasks required to maintain a particular function's existence in the city. The more unique and special an economic function, the higher the threshold.
- **Range of good or service** - the maximum distance a person is willing to travel to obtain a good or service. And, the maximum distance at which a good or service may be accessed. How far is a consumer willing to travel? Central place theory assumes consumers will not be willing to travel as far for lower central place functions. What if the consumers may no longer need to travel because the transportation system now delivers?
- **Spatial competition** - central places compete with each other for customers. Central place theory assumes central paces will be located farther away from each other, because consumers are more likely and willing to travel a longer distance to obtain higher central place functions.

2.5.1 Circular city naming

A.k.a., 2D Circular grid, radial grid/plot, polar grid/plot, hemisphere mesh, circular layout, circular or polar graph.

A circular city has the following identifiable elements (Hakan, 2020):

1. **Circle** - A circle is the path traced out by a point, moving in a plane, that is always a fixed distance (the radius) from a fixed point (the centre).
2. **Center** – Location of the grid origin, [point]
3. **Central area** - area of central most circle. The formula for the area of a circle is:
   - $A=\pi r^2$
   - Wherein, A = Area; π = 3.14...; r = radius
4. **Inner** – The inner radius of the grid, [number]. Radius of inner circular sector.
5. **Outer** – Radius of the grid, [number]. Radius of outer circular sector (radius of city boundary).
6. **Sectors** – Sets the number of sector dividers, [array]
7. **Rings** (belts, circulars) – Number of concentric dividers of the grid, [number]

The parts of a circle are (*The circle*, 2011):

1. **Radius** - A radius is any interval (or line segment) drawn from the centre of a given circle to any point on the circle is called a radius, (plural radii).
2. **Diameter** - A diameter is any interval joining two points on the circle and passing through the centre is called the diameter of the circle.
3. **Semicircle** - A diameter divides the circle into two congruent parts. Each part is called a semicircle (2 total semicircles to a whole circle).
4. **Quadrants** - If a radius is drawn perpendicular to the diameter in a semicircle, there are two congruent quadrants (4 total quadrants to a whole circle).
5. **Sector** - Any two radii divide the circle into two pieces. Each piece is called a sector (from the Latin word secāre − to cut).
6. **Circumference** - The distance around a circle.

There are a variety of different layouts of circular city, including but not limited to:

1. **Concentric** - denoting circles, arcs, other shapes that share the same (com-) center.
2. **Overlapping** - denoting circles, ars, other shapes that overlap.

2.5.2 City layouts in community

A.k.a., City shape, urban pattern.

Most cities in community are of a circular arrangement (a.k.a., radial-concentric, ring-radial, circular radio-centric, or polar coordinate configuration) with the central area acting as a representative centerpiece of that particular city. There are non-circular cities, some of which are non-circular because the geography won't allow for a fully circular configuration. Cities aren't generally built on a flat surface, even planned cities have to work around natural features in the terrain; that is, to the degree to which the site has been appropriately selected and the terrain is capable of being modified. The circular city is simply a theoretically "optimal" design, local topography and geography will, in many cases, change the design slightly.

NOTE: *Living organisms have bi-lateral symmetry. If the city is viewed as a living organism, then it may be designed with bi-lateral symmetry (i.e., city symmetry).*

The proposed circular configuration of many of the cities in community is not a just stylized architectural conceptualization. It is the result of reasoning and evidence into providing an environment that can best serve the needs of the inhabitants and conserve resources. The circular arrangement effectively permits the most sophisticated use of available resources and construction techniques with minimum expenditure of energy. The efficiency of the circular design allows us to make available to all people the most advanced amenities that our knowledge and energy can provide.

A circular city is most practically divided via pathways into areas known as [radial] sectors and circular belts (a.k.a. "circulars" or "rings"). The radial sectors (separated by pathways) are subdivided by circular belts (also separated by pathways), which extend outward from a central point, forming a widening circular grid structure. As the circle widens, more circular belts follow until the perimeter is reached wherein the environment is allowed to return to wild nature without any form of sprawl. In other words, these circular cities are composed of a central area beyond which the geometry takes the form of radial sectors and circular segments. In most configurations, there is a differentiation of primary functioning between belts (and sometimes within segments of a belt itself). In other words, each circular belt (and/or radial segment) maintains a particular set of functions, some of which will be unique to that circular belt and will give the belt its name. Other functions are shared between belts. The core function of the recreation belt, for example, is to provide recreational services and structures. Secondarily, however, the recreational belt maintains permacultural land and aquatic spaces for the growth of food and natural beauty. Although every circular belt will have a core identifying function, all belts are multi-functional.

There are a variety of reasons why a circular city scheme is more efficient than other city layouts. Firstly, when you start at one point on a circle, and move along that point, you eventually come back to the same point. When it's a linear city within which you are moving, you have to travel back again (i.e., backtrack) over the same area [instead of just going around). Hence, when traveling within a circular city someone could easily return to the same place from where they started without having to take the same route back, as is the case with most linear cities. Secondly, circular designs place frequently used facilities (mass transit, medical, and other common access locations) near the center. This puts most of the residential population very near (in time and space) to the city center, and ensures that travel throughout the city is relatively easy. Hence, no matter where you are in a circular city, you would be within a reasonable distance to access every facility the city has to offer. A circular shaped city ensures that no [access] point on the circle is ever further away than half the circumference of the circle itself, which is an important design consideration for emergency response. Conversely, a squared shape maintains that no point is further from another than the "Manhattan-distance" (i.e., the distance between two points, as 90° horizontal and vertical paths on a square grid; versus an acute diagonal(s) with a circular grid). Fourth, a planned circular design minimizes the length of all transportation and distribution lines (in comparison to a linear design) -- less to build, less to maintain, and hence, more efficient. Fifth, consider that a grid inside a circle would combine the advantages of best use of space with a most understandable addressing system. Of course, either a square grid or circular grid are better than a random or disorganized configuration. A circle, however, provides the most efficient form of infrastructural elements required for its outside perimeter. Only 1 shape of interlocking element is required over 2 shapes (straight and right angled) for a square. Sixth, the circular design allows for one "pie-like" sector of the city to be designed, and then replicated around the circle six to eight times (with slight adaptations for functional differentiation) to form the entire city. In the design and production of a circular city we work out 1/6th or 1/8th (for example) of the city system, and then we reproduce it around a central point. The replication of a radial sector around a central axis (returning to the original sector itself) uses fewer resources than conventional construction methods for linear cities. In market terminology, these cities are extremely cost efficient because only one radial sector needs to be designed, which can then be duplicated repeatedly and slightly versioned for the completion of an entire city. Seventh, a circular layout is easily replicated at different scales. These cities can be designed for a couple hundred people, or scaled up to population sizes of 100,000 or more. And finally, at least for this discussion, the circular arrangement is also a useful geometric design for mirroring natural symbiotic cultivation cycles. Circular symbiotic farming, for example, is often applied as part of the last circular belt of these cities.

In general, a well-designed and aesthetic circular city tends to feel more harmonious and open than its equivalent as a linear city. We do live on sphere (of sorts), and from a two dimensional perspective the planet upon which we live takes the shape of a circle. It may be further interesting to consider that our eyes, the stars in the sky, including our sun, and the moon are also all circular in shape. Even our galaxy has a circular symmetry. It may be interesting to consider that the motions of nature move in spheres and rings, and all cosmic bodies seem to move in spiralling arcs.

The round architectural shape of a circle provides a natural sense of unity. Historically, it has also been a practical form of defense against dangers coming in from all sides. Further, the circle is one of the emblematic tools to express one-ness in a visible environment. Corners break the non-hierarchy of unity. All planets and suns take the shape of a circle. There is clearly a connection between central planning (optimal socio-technical organization), the search for unity (optimal information organization, and the shape of a circle (optimal structural/geometric symmetry). (Delen, 2016)

The relative current scarcity of circular cities on the planet in the early 21st century can probably be explained by considering two causes. Firstly, the prerequisite for a circular city is a suitable geography (the natural factor) and a deliberate plan to continue city development along concentric lines (the human factor). Ideally, the round city is situated on a plane (or terrain modified environment) without significant natural obstructions. Such natural areas are common all over the world, but there are, nevertheless, very few circular cities that are built in those natural, ideal geographical areas. Hence, the absence of circular cities on a wider scale must have another reason, which is likely found in the human factor. It seems that people in power are not interested in the idea of circularity and its inherent neutrality. The combination of a strong government, which can implement ideas by force versus a non-hierarchical message of the circle-in-general is an dissonant one. Powerful governments, based on a vigorous application of law and order, are hardly ever the keepers of peaceful ideas. The main reason is, that otherwise they would not be in command. (Delen, 2016)

> *"The preference of the circle as an architectural feature is the result of a resistance and opposition against squareness. However, the circle is also – in a non-oppositional ambiance – the beginning (or end) of a path of insight." (Delen, 2016)*

The growth of most cities in the 21st century is the result of ad-hoc market and political decisions. The concentric design needs a deliberate planning in a fairly unprejudiced setting. Rapid urban developments have no time for the relative forethought of a circular configuration. Such types of city layouts only come into being under special (i.e., thoughtful) circumstances. (Delen, 2016)

Continuum approximation (CA) optimization models can be formulated and tested to design an optimal city-wide transit system with correlation to optimal city layout. Chen (2015) used two models for comparison. Model 1 assumes that the city streets are laid out in ring-radial fashion. Model 2 assumes that the city streets form a square grid. Therein, Chen et al. assumed transit routes lie atop a city's street network. Model 1 allows the service frequency and the route spacing at a location to vary arbitrarily with the location's distance from the center. Model 2 also allows such variation but in the periphery only. Chen et al. shows how to solve these CA optimization problems numerically, and how the numerical results can be used to design actual systems. The results show that Model 1 is distinguished from Model 2 in that the former produces in all cases: (i) a much smaller central district, and (ii) a high frequency circular line on the outer edge of that central district. Parametric tests with all the scenarios further show that Model 1 is consistently more favorable to transit than Model 2. And, cost differences between the two designs are typically between 9% and 13%, but can top 21.5%. This is attributed to the manner in which ring-radial networks naturally concentrate passenger's shortest paths, and to the economies of demand concentration that transit exhibits. Thus, it appears that ring-radial street networks are better for transit than grids. (Chen, 2015)

2.6 City expansion

> **INSIGHT:** *What is a tumor? A tumor is a growth untethered to the consequence of it growing; a growth for its own sake, otherwise known as a suburb. A suburb is a type of societal tumor.*

Individual cities can grow in three ways:

1. **Outward** - expanding horizontally.
2. **Upward** - expanding vertically.
3. **Toward greater density** (a.k.a., densification, in-fill)- expanding interstitially (i.e., filling up every free space and reducing the space available for inhabitants); often the least pleasant for inhabitants.

The market's solution to overcoming population congestions is, most often, to spread out horizontally. All early 21st century cities have done this (i.e., spread and sprawled outward), only to create more problems. Moreover, expansion is generally not uniform, making the problem of transportation even more complicated. This has come to be known as urban and sub-urban sprawl.

Cities in community are designed with a planned, specific carrying capacity. When a city hits a certain size, it stops and mostly thereafter, everything is allowed to return back to nature between this and the next city; there is no urban sprawl. The iterative design for a city in community is "organic" to the extent that new information evolves the system; but, its operation is planned, and so, there is no sprawl (i.e., no "suburb") or haphazard/chaotic development (as is the case with the "organic" development of nearly all prior cities).

> **INSIGHT:** *To "suburbia" a society leads to the separation of the individual from a place of meaningful effort, meaningful relationships, and meaningful results. Do you live in a suburb? Is it considered acceptable to randomly hug your neighbour?*

Cities in community are not meant to be ever expanding, as is the case with early 21st century cities and suburbs. Instead, circular cities can be reconfigured internally, but the diameter is mostly fixed. Instead of expanding cities horizontally (i.e., over the surface medium they are built upon), a new city is created nearby and connected by a transportation network so that nature is left between cities. Community cities can be iterated and updated internally, and also expanded vertically, but they are not intended to be expanded in surface area coverage.

It is true that squares can be more easily compacted [next to one another] than circles, but when designing city systems for community, beyond the perimeter of the city, the environment is allowed to return to wild [caretaken] nature. So, whereas a linear or squared city would just continue to add more blocks/modules [to itself]; instead, community would allow a return to nature prior to the creation of another [circular] city. Note that the one exception to this rule may be extreme desert environments where there is little to no life beyond the perimeter of the city.

> **NOTE:** *In community, we don't want indefinite [city, economic, or otherwise] expansion on our finite planet. In general, when a city reaches carrying capacity, another city will be built, separated by nature some calculated distance away from the prior.*

A city with square blocks can expand indefinitely by placing another block next to the prior, while a city with a single circular block cannot do so with compact geometric alignment. A circular city is one circular grid reducing to a central axis. Of course, if a circular city requires expansion for some reason, it is still possible to do so with geometric alignment by extending the city radially, segment by segment.

2.7 Circular city assembly

There are two basic ways to assemble a circular city (note that these two ways can be mixed):

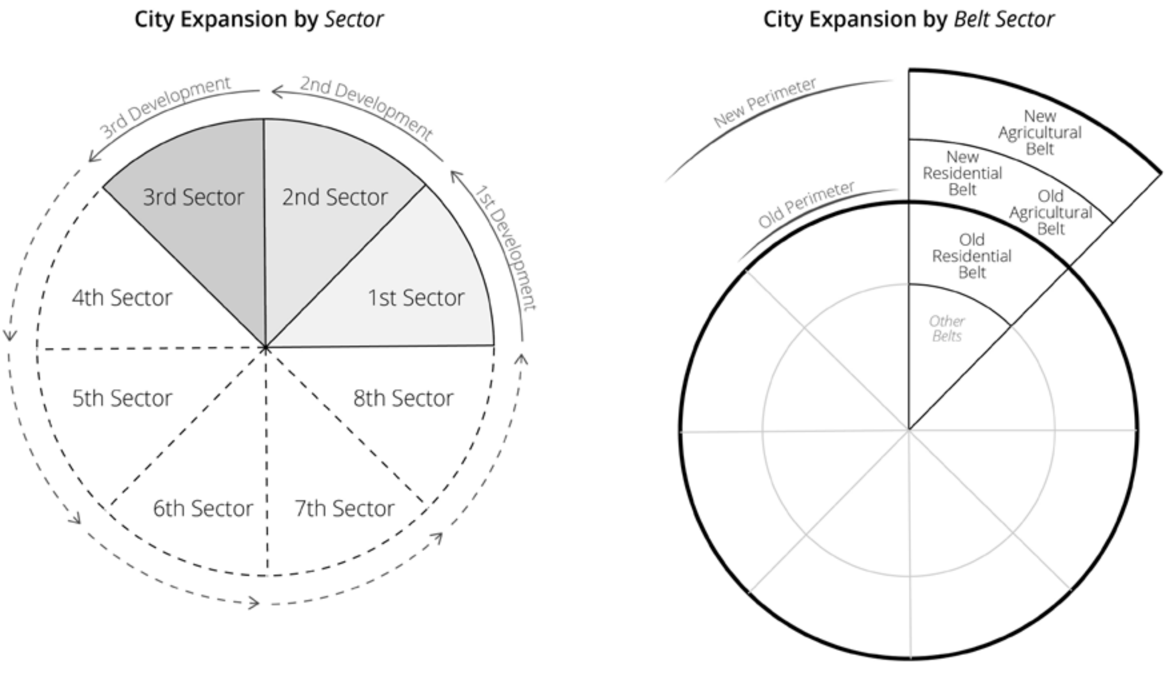
1. Radially - radial segment by radial segment.
2. Circularly - circular belt by circular belt until the planned perimeter is met.

Note that if circular farming was used on the outer segmented belt during the city's phased construction, the soil base could be built up as the city was assembled (belt by belt) to its planned size. For example, originally the city may have only have three circular belts constructed with planned eight. The third circular belt of the initial construction could a circular farm, which would build up a soil base on that belt. When the next belt is added, the circular farming is moved to the fourth belt, and so on. This would obviously create a more lengthy time frame for the construction of said city, but the result would be a higher quality soil base for all belts where circular symbiotic farming was applied.

2.8 Computational and mathematical modeling for cities

Data acquisition and 3D modeling has enabled the dynamic modeling of physical phenomena, objects, and human behavior, including the simulation of complex environments and problems. Computational models can be applied to almost every aspect of city design, decisioning, and operation. Therein, computational models require relevant data and are used to produce indicators for decision support by each subsystem of a habitat service system. Additionally, computational

Figure 6. *Figure on left shows city expansion going sector by sector over time. Figure on right shows sector-ring by sector-ring over time.*



models and their associated software are necessary for optimization, improved design and decisioning. (Stjepandic, 2019)

Computational modeling is fundamentally related to mathematical modeling. A mathematical mode of a the human domain is a formal representation of individuals' attributes and/or desired requirements of a design. For instance, the definition of values and their encoding (transformation) into decisioning is significant for design. Mathematical modeling allows for precision, which can be used to achieve traceability, robustness, certainty, and better rapport with reality. Further, a mathematical model of the functional domains of a habitat service system would represent those functions formally. A formal representation of functions is a prerequisite for representing functions in computers. A mathematical model of the behaviors represents formally the behavior of the physical model. Wherein, the mathematical model of the physical domain is a formal representation of physical variables, design principles and physical principles of the design of a city. (Stjepandic, 2019)

INSIGHT: *We must operate within the carrying capacity of the city as we do similarly within the limits of the earth itself.*

2.9 Evolution and appropriate habitat design

There exist three general evolutionarily-oriented principles for habitat [re-]design:

1. Constant habitat features: If there was some habitat feature that was constant during all of a species evolution, then it must be accounted for in the habitat's design. Take gravity for example, humanity has never known life any other way. If gravity is altered significantly, or it isn't there at all, then the organism will start to face some very serious issues, such as bone density and functional strength - this is seen in space.
2. Cyclical habitat features: Some habitat features are cyclical, like night and day. In the case of a cyclically variable habitat feature it is possible to modify it with a degree of deviation and to the extent that the organism is adapted to it changing. But, if it gets completely out of cycle, out of sync, or becomes monotonous on one side or the other, then the organism is likely to experience dis-ease - this is seen on submarines and with shift-work (where shift workers experience high rates of cancer).
3. Variable habitat features: If a species evolved for much of its evolutionary history with a variable habitat feature [within some bounds], it is probably adapted to that feature remaining varied. So if you have temperature with a dynamic fluctuation through the seasons, or through the days/nights, or maybe the surfaces that the organism moves on varies, and things become too monotonous, then the organism is likely to start experiencing dis-regulation. Variation can be healthy and present adaptive advantages. For example, the night environment is colder than the day environment, which facilitates sleep onset and a lack of waking during the night.

The human body is dysfunctional in certain environments. It is not adapted to certain environmental structures, and it maladaptive to others. And, without knowledge and a realization of the environmental territory into which one is entering, or being conditioned, an individual could kill oneself or cause serious harm to the continued optimal functioning of its organism if it doesn't account for its principle redesign (qualified by a hermetic stress response to prevent fragility).

3 The life radius

A.k.a., The human life movement space.

A city is essentially a demarcated material ‘life radius’ within which a population sustainably controls environmental variables and optimize human fulfillment. Individuals spend the majority of their time in the same places, and that environment dictates how easy or difficult it is to make healthy life choices and express one’s highest potentials. The term “life radius”, itself, describes the space where a population spends the vast majority of their lives (~80 - 90%). To clarify exactly what a life radius is, someone might ask themselves, What are the places I walk to and through on a daily basis? In this life radius are the spaces and places frequented on a regular (e.g., daily/weekly) basis.

> **CLARIFICATION:** *A ‘life radius’ is a place where individuals spend approximately 90% of their current life.*

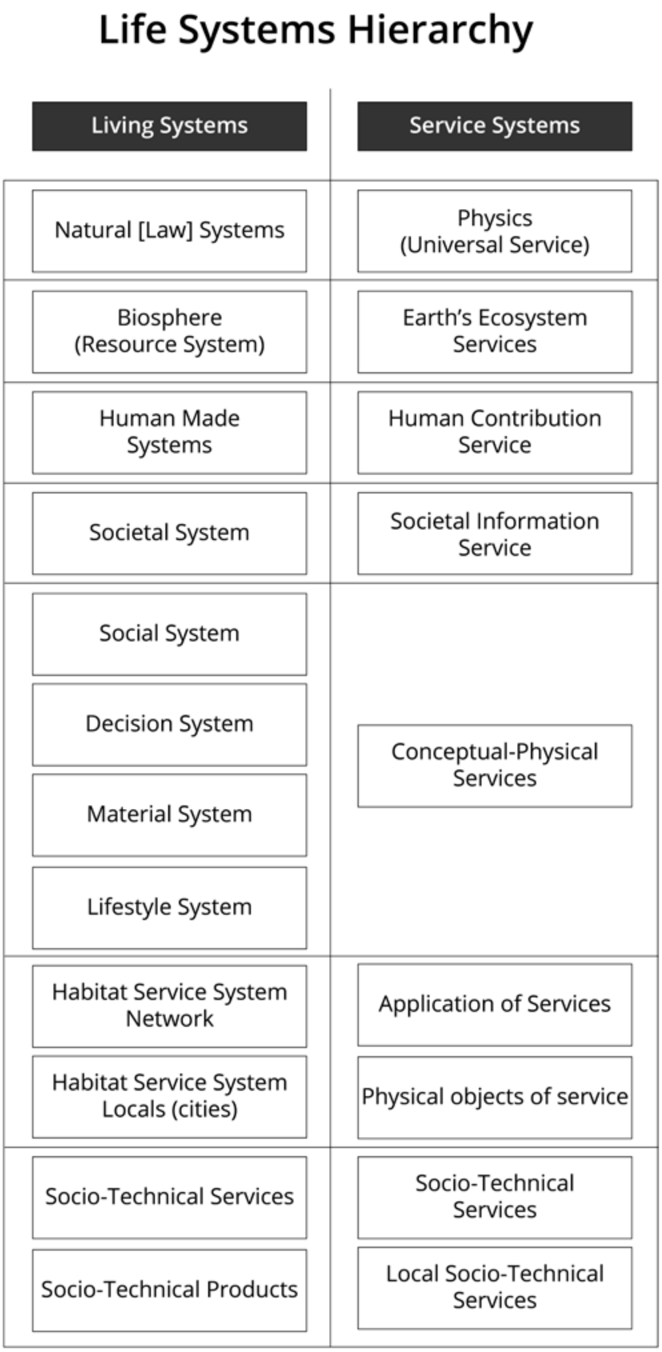
Everything that occurs within the life radius is considered to have an impact on everything else, making it possible for an aware population to control and optimize for their fulfillment within that life radius.. When individuals have to drive a car, that radius can be quite large. But, the ideal life radius is much smaller than city arrangements where cars are necessary. In community, cities are designed at a scale based upon the human being, and not the motorcar or some abstraction. To clarify exactly what a life radius is, someone might ask themselves, What are the places I walk to and through on a daily basis?

> **NOTE:** *Individuals [are likely to] entrain to their environment. If individuals live in a depressed environment, they are likely to be depressed (or, become desensitized to the depression). If individuals live in a happy environment they are likely to be happy, and become sensitive to the happiness of those around us.*

Community is designed in a people-oriented way. The average human being walks two kilometers in approximately twenty minutes. What if that two kilometer walk was beautiful, attractive, safe, enjoyable, and an individuals could meet their needs, contribute, and develop themselves, with others who are doing similarly. A bicycle extends the radius, or makes movement in the radius more efficient. Certainly, a bicycle or mass rapid transport system has a potential of extending what may otherwise be the ideal walking life radius. But, the point is that “you” want most of the things “you” are going to do, for some large percentage of “your” time, to be inside that radius. Having access to what is needed within a walkable radius is strongly correlated with well-being (happiness).

Think about your own life for a moment, where do you work, where are your friend’s homes, your enriched

Figure 7. *The societal life systems hierarchy of a community-type society. The left column contains systems that are dynamic and feed back into a total human life system. The service systems in the right column provide the informational and physical generations (material relationships) that complete human material requirements.*



gathering and relaxation spaces, and the locations that produce and distribute your material necessities? Of those key things that compose your life radius, how many can you access by foot or bicycle, and is the experience safe, comfortable, and enjoyable.

> **INSIGHT:** *It is possible to make the healthy and "right" choices the easy ones, with appropriate challenge and preference layered in.*

In community, the life radius is designed to:

- Generate a social and economic decision structure, an environment, where it is it easier to get up and move, eat healthy, make new friends, find a reason for being, and live longer, more optimized lives.
- Create an environment where people move naturally each day without thinking about it. Community makes it pleasant and enjoyable to leave ones dwelling and participate in activities.
- Facilitate healthy food choices while bringing attention to foods that are more nutritious (and hence, flavorful).
- Support personal interconnectivity—between individuals and community activities, teams, and groups.

> **INSIGHT:** *Historically, cities grew because more people move to them than died inside of them.*

3.1 Moveability / walkability

The more thought responsive the world becomes (due to technological automation), the less individuals technically have to move their bodies. And so, humanity might as well design its city environments so movement is intrinsic and facilitated.

3.1.1 Needs versus inculcated expectations

In concern to city design, the question must be asked, Do we need to drive anywhere in the city? Certainly a city population has a need for transportation (personal locomotion, mass locomotion, and emergency locomotion), but is there a need for transportation via cars within the city boundary. It is possible to design and plan a city environment so no one needs a car. It is possible to create walking garden cities where walking and biking are the primary form of movement, and where vehicles are used for emergencies, mass rapid transportation, scenic transportation, and automated distribution functions (e.g., delivery robots).

4 An example integrated city system

Generally speaking, at the level of the material architecture of a human community with a sufficiently large population, and access to digital information technology, are circularly configured walking-garden cities. As we zoom out from one of these cities we see a branching network of cities, each separated by nature. Different cities in the network may display different functional configurations and architectural aesthetics, although they are all still based around a unified community information system. While many of the cities in the network would be circular, others may be linear, underground, or constructed as floating cities in the sea.

This example will first start with a description of the center of the city and work my way outward through the different circular belts. Take note that the stylized elements of buildings and areas in these cities can be customized to the preferred and traditional cultural aesthetics of the local geographic population. For example, buildings in a community-city in China, Japan, India, Europe, the Americas, Africa, or the Middle East may have stylized design elements traditional to those locales.

The following is a hypothetical example. Herein, the land area belts of the circular city are operationalized under the service of a habitat system for functional differentiation. Each belt is a spatial boundary allocated to a different functional service. Between the belts there are circular pathways, and positioned radially around the circle are radial pathways.

4.1 The central area

The first area of the circular city arrangement I would like to point out is the city's center; its central access point. Here in the center of one of these circular cities you may find medical care, conference centers, exhibition and art centers, and a whole host of other spaces where social interaction occurs. This central area may also be a transportation hub if the city includes a mass rapid transportation system. Note that if medical facilities are placed in the central hub, then you are never further away from receiving medical care than if you were in the same belt in another sector of the city, which is an important consideration for an active and playful population. And of course, under other city configurations the central area may not have any buildings, but instead it may be a garden for common gathering and natural beauty.

4.2 Permacultural gardens

Moving out from the central area, this configuration [we are imagining] has permacultural and aquacultural walking gardens and parks. These are beautiful landscapes organized for food cultivation and aesthetic relaxation. As you walk through them fresh food is available seasonally for harvest, and there is ground for playing and contemplation.

4.3 The habitat systems service sector (InterSystems Operations Sector)

The next circular belt out is mostly composed of buildings used for the completion of work relevant to the continuity of the entire city system (it is more commonly known as the InterSystems Operations Sector). These buildings house access hubs, maintenance and operations facilities, as well as research and production spaces. Here, we primarily complete work which updates and cycles services and technologies through the city. All belts are multi-functional, and so within these buildings there are also many common access spaces for a wide variety of technical- and creativity-oriented activities.

4.4 Recreational area

As we move away from the service belt we come to the recreational area, which has courts, gyms, and all of the games and recreational activities that people require, amongst beautiful terrain and landscaping. This belt has art centers, theatres, and various spaces for practice and entertainment. There may also dining facilities here, and other amenities.

4.5 Low-density house dwelling area

As we move outward, again, we come to the low-density dwelling and housing area where there are winding streams, ponds, waterfalls, and lovely gardens throughout, giving each dwelling a view of beauty and a feeling of being at restorative peace with the world. The residential area of the city continues the idea of coexisting harmoniously with nature. All of the houses are similar in their modern rounded design, but at the same time are very different. Their uniqueness is a reflection of the owner's personality and desired functioning of the home. The architectural elements of all dwellings are flexible and coherently arranged to best serve individual preference. The features of all dwellings in the city are selected by the occupants themselves.

In between every home are natural barriers like bushes and trees, isolating one from another with lush landscaping. So, people who prefer to live in houses and maintain gardens may prefer to live in this area.

4.6 High-density dwelling

The next belt we come to primarily functions for high-density dwelling. Its dwellings are for those who prefer apartments. The reason some people may want to live in an apartment is because the apartment buildings themselves have a large number of services built into the tower, providing immediate and close access for those who might want that sort of dwelling placement. People who choose to live in apartments may prefer a more socially dense dwelling arrangement. These dwellings are also above the ground, and so, they provide beautiful views of the city and the surrounding natural environment.

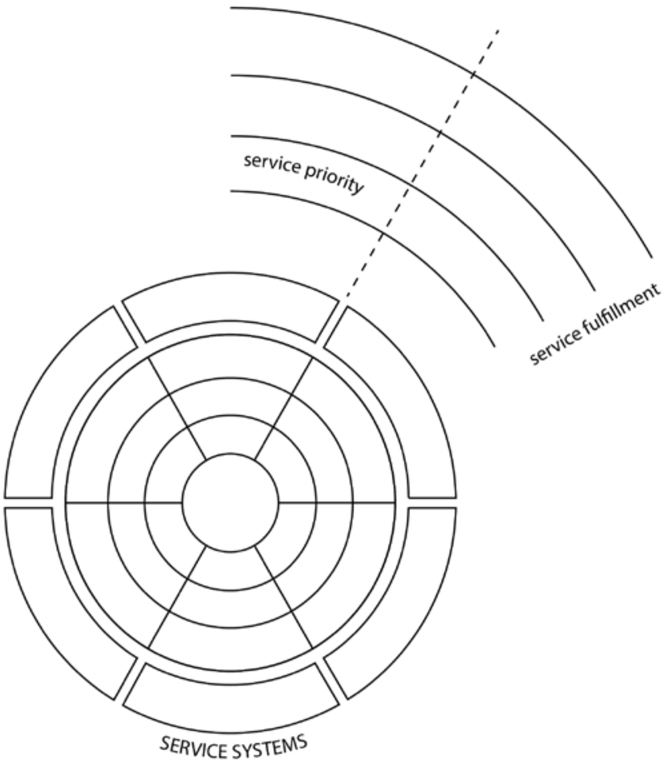
Secondarily, this belt maintains energy production systems, as well as lovely gardens and relaxed common gathering areas.

4.7 Water channels and controlled cultivation

Passing out of the high-density dwelling belt on our way to the outer ring of the city we come to the primary food cultivation belt in-between two water channels. On the food cultivation belt we organically grow a wide-variety of plant and insect species, both outdoor and inside greenhouses. Here, a beautiful walking and bicycling path encircling the entire belt. The primary function of this cultivation belt is to grow sufficient food for all the inhabitants of the city.

When looking at the water channels consider for a moment the wisdom of our ancestors in their choice to developed their living systems around a water source. Here, the waterways provide water storage, harvesting,

Figure 8. *Concept diagram depicting service systems with service priority (i.e., some services are prioritized) and service fulfillment (accountable degree of completion of service system demand).*



irrigation, and purification. On the water channels there are water harvesting atmospheric generators with solar distillation units. These evaporative condensation systems are one means by which the city creates clean drinking water. And, at least one channel is always available for swimming.

4.8 A natural barrier

Just beyond the final waterway is a ring constructed as a geomorphic vegetation-barrier. It is designed to prevent ecological disruption to the inner city and purify environmental run-off from the next belt outward. The vegetation selected for this natural barrier will have a second purpose, it will be used for harvesting into food, textiles, and many other useful materials.

4.9 A circular farming system

In this configuration the outer perimeter ring is [in part] a “circular farm”, a holistically planned grazing system also known by the names circular symbiotic cultivation, regenerative agriculture, rotational grazing, and syntropy farming. It is a biomimicry process that mirrors what occurs in nature. Here, the “farming” follows natural ecological cycles. This circular area is primarily a combination of pasture and orchard land that we move different animals through in a particular order to mimic natural cycles, which builds our soil base and provides food.

In this area there is grass between trees, and often, when left unchecked, the grass will grow up and choke out the tress (same with shrubs). Early 21st century society generally prevents this consequence by using a lawn mower. But, nature provides an alternative. Imagine running a number of different organisms around this circular ringed area. We send cattle through the orchard and let them mow down all the grass. And, as they go the cattle fertilize the tress. They deposit their waste, and then, trample it into the ground to create fertile, carbon rich soil. A few days after the cattle, we send the goats, who eat the shrubbery that the cattle wouldn’t necessarily eat. The goats also climb up and prune the bottom 6 feet of the trees. They also fertilize. Pigs are run through as left-over waste consumers. Then we send through the chickens in a mobile chicken coup. The chickens also fertilize the soil and eat all the bugs that hatch from the manure of the first two ruminants that went through. Chickens come in after the pigs have dug up big clumps of grass. They “cleaning out” the area and fertilize with their high nitrogen manure. So, at the least, we intentionally run 4 different animal species through this area, and as a result, we get multiple cultivations, we build up our soil base, and we have the opportunity to play a role in the well-being of other symbiotic species, while giving ourselves a picturesque environment to enjoy in a variety of fashions.

Among the circular farm, this ring may also be used for recreational activities such as biking, golfing, hiking and riding. Areas herein may be set aside for renewable, clean sources of energy, such as wind, solar, heat concentrating systems, geothermal, and others. There may also be large activity domes positioned around this ring if that is what the population of a particular city desires. Further, there could be lower-rise apartment type structures close to the outer edge for people who prefer apartments, but would like a more outdoors type of living, close to where the city returns to wild nature. And finally, this outer perimeter could be considered another natural barrier, designed to prevent ecological disruption to the inner city.

4.10 Return to nature with care

Beyond the outer belt we allow the environment to return to nature, while still caretaking our total habitat. When a city reaches its planned size, we stop, and let everything go back to nature between this and the next city. There is no urban sprawl; mostly, we let everything return to nature between cities -- we let the environment return to its natural homeodynamic equilibrium. Out in nature we can wild food forage and re-learn the skills or our ancestors. Here, we ask ourselves, “What is it like to be just another animal in the wild?”

4.11 Transportation

In concern to transportation, these cities generally contain two to four primary transportation gateways (i.e., entrances and exits). Few transportation gateways are needed for the city because of its efficient design. Transportation within the city and between cities is shared between autonomous transveyors, specialized electric motor vehicles, self-powered vehicles (e.g., bicycle), and mass rapid transporters (MRTs) – all in the form of emissions-free transport. The design of these cities removes the need for each individual (or family) to have a personal automobile. Of course, mostly, these cities are designed for walking. Some cities, however, are large enough to necessitate transveyors and/or an MRT system within their limits.

I would like to leave you with one final thought: With a population of over 7 billion people on the planet it is essential for us to merge our knowledge of nature with a fulfillment-orientation that can guide the things we do and the cities we create.

5 Biophilic design

Biophilic design is the practice of connecting people and nature within our built environments and communities. The promising new field of biophilia suggests that human beings have evolved with certain basic aesthetic and physiological needs: the presence of vegetation, water, sunlight, animals, and also the geometric relationships that have accompanied our evolutionary experiences with these structures. By tapping into this rich vocabulary of biophilic design elements, we can have an extremely rich variety of design possibilities — a rich range of artistic expression — while still meeting the needs of human beings. And within the same life-affirming process, we can meet the ecological needs of the environment too.

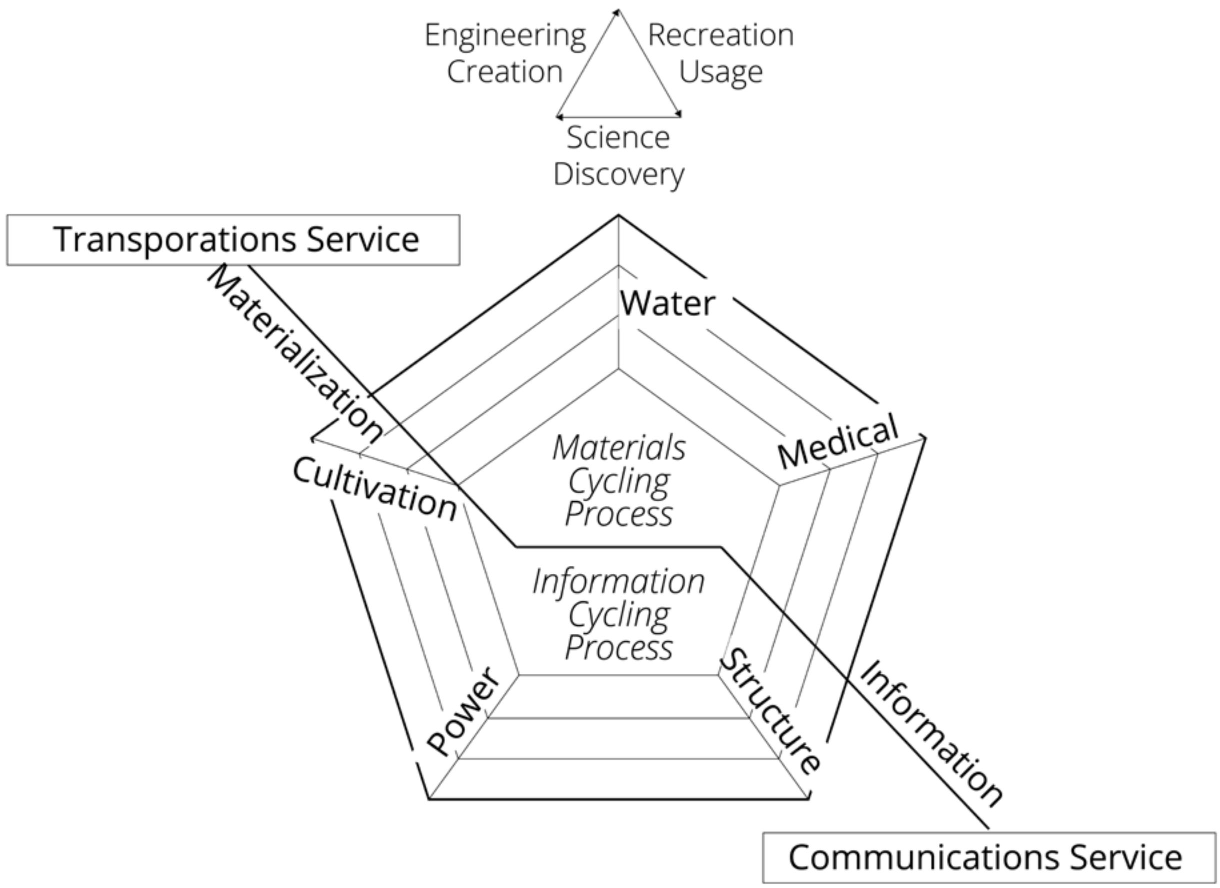
True, people have enormous varieties of experiences and tastes — and it's wonderful that they do — but these phenomena are generated by a common set of structural processes that are identifiable and shareable. Some experiences are unquestionably damaging to health and wellbeing, in the same way that, say, the structure of car exhaust molecules is damaging to health and wellbeing. It does no good to say our narrative about car exhaust is such and such, we want people to experience it and be provoked by it — that will not change the fact that we are making people unwell.

Doctors have learned that certain aspects of the patient environment promote well-being, and they now use this "evidence-based design" to improve the quality of life of their patients. In the same way, adaptive, human-scale architecture and urbanism rely upon discoverable rules of design. We proposed the existence of such rules while at the same time conjecturing that a non-adaptive aesthetic is easily reached from the adaptive design rules by simply reversing them. That is, since guidelines for designing adaptive, contextual environments are known instinctively, do the opposite to generate a form that strikes an observer by its visual novelty and lack of context.

> **INSIGHT:** *Awareness of a timeless language is present in people, but they learn to suppress it.*

Nature sees everything as a complex and continuous interaction where information is not separate from matter. Information has materiality, which has

Figure 9. *Integration of life support and technological support into a model that produces the likelihood of societal structures of the informational and spatial order that sustain, and may even optimize, human fulfillment under dynamic and changing environmental conditions.*



properties.'Digital materialization' refers to a two way conversion between matter and information. A system with the following four characteristics:

1. Symbolic - it has to be similar to the way we deal with other exact systems, like mathematics, or indeed, is mathematics itself.
2. Volumetric - it can't just be defining a 2.5D (a set of surfaces arranged in 3D space) it has to define at least 3D space, if not 4D.
3. Constructive - it needs to be modular; constructors must be able to work in chunks (e.g., like a Legos construction).
4. Continuous - it has to have continuous/infinite surface. can keep zooming in, you can keep going down to any resolution you want, you aren't limited.
5. Exact[ness] - the system has to have exact inputs and exact outputs.

6 Aesthetics

> **INSIGHT:** *There are places and environments that are just going to depress your heart rate variability.*

The human eye "likes things" in certain positions; it finds some positioning and proportioning more pleasing. Hence, community environments are often built in the proportion(s) of true beauty. Community developers often seek to create a sense of harmony and alignment with the pattered expression of nature in all space.

> *"The bad formation of towns influence the bad formation of minds." (Pemberton, 1854)*

How the body relates to a space can be studied independently of what is going on in the mind (e.g., ergonomics), but how the mind engages space has to include the body and the brain of the individual. At the level of core, or basic, consciousness, humans are consciously and unconsciously registering the environmental variables' effects on their nervous system -- heat, light, noise, smells, tactile sensations, and a perception of movement and spatial orientation arising from stimuli within the body itself. All of these sensations are silently registering in the viscera as well as the somatosensory cortex via signals of which individuals are often not aware. At the level of extended consciousness, individuals are simultaneously experiencing space as assembled by their sensory system and combining this experience with memories of places similar to the one they are in. Individuals' minds are sorting through all of this to let them know they are dealing with a "reality". Part of the brain's internal environment is generated by a ceaseless pressure to seek out new stimuli. This is why humans are sometimes called "infovores", a term coined by neuroscientists Irving Biederman and Edward Vessel to mean a person who desires and seeks out information gathering. This hunger for information is one of the fundamental properties of the brain, and it is reflected in human individuals' most basic reactions (Biederman, 2006).

By understanding the biological basis for stress, we understand the potential for induced illness within a cognitive environment, as well as how to induce wellness. By understanding how lighting, acoustics, thermal conditions, and windows affect cognitive activity, we will have evidence for enriching the environment.

> **NOTE:** *In Japan there is a term for the rejuvenation provided by being in nature. Shinrin-yoku, also known as "forest bathing" is a simple practice for enhancing health through sensory immersion in forests and other naturally healing environments.*

Parks and other green spaces make people happier, and the proof is in their brain activity. The understanding that nature-based environments facilitate health is the reason that walking in nature or a park is recommended for rehab patients and athletes in recovery. Nature also, perhaps, facilitates resource appreciation. Certainly, survival training facilitates an appreciation of resources.

> **INSIGHT:** *Our environment has a profound impact on the way we feel and perform. We can create through cooperation and intelligence a micro environment inside nature that is supportive of us and our evolution.*

Scholarly references

- Aspinall, P., Mavros, P., Coyne, R., Roe, J. (2015). *The urban brain: analysing outdoor physical activity with mobile EEG*. British Journal of Sports Medicine, 49(4), pp272-6. DOI:10.1136/bjsports-2012-091877 [ncbi.nlm.nih.gov]
- Babar, M. and Arif, F. (2017). *Smart urban planning using big data analytics to contend with the interoperability in internet of things*. Future Generation Computer Systems, 77:65–76.
- Batten, D.F., (1995). *Network cities: Creative Urban Agglomerations for the 21st Century*. Sage Journals. DOI:10.1080/00420989550013103
- Batty, M., Cheshire, J. (2011). *Cities as flows, cities of flows*. Environment and Planning B: Planning and Design 2011, volume 38, pages 195-196. DOI: 10.1068/b3802ed [people.iup.edu]
- Bermudez, J., Krizaj, D., Lipschitz, D.L., et al. (2017). *Externally-induced meditative states: an exploratory fMRI study of architects' responses to contemplative architecture*. Frontiers of Architecture Research, 6(2), pp123-136. DOI: 10.1016/j.foar.2017.02.002 [sciencedirect.com]
- Ben-Joseph, E., Gordon, D., (2000). *Hexagonal Planning in Theory and Practice*. Journal of Urban Design, 5(3), pp237-265. [web.mit.edu]
- Biederman, I., Vessel, E., (2006). *Perceptual pleasure and the brain: A novel theory explains why the brain craves information and seeks it through the senses*. American Scientist. [cvcl.mit.edu]
- Cats, O., Vermeulen, A. (2019). *Modelling Growth Principles of Metropolitan Public Transport Networks*. Department of Transport & Planning, Delft University of Technology. [transp-or.epfl.ch]
- Chen, H., Gu, W., Cassidy, M.J., Daganzo, C.F. (2015). *Optimal transit service atop ring-radial and grid street networks: A continuum approximation design method and comparisons*. Transportation Research Part B: Methodological, 81, 755–774. doi:10.1016/j.trb.2015.06.012
- Chen, Y., Nakazawa, J., Yonezawa, T., Kawsaki, T., and Tokuda, H. (2016). *Cruisers: A public automotive sensing platform for smart cities. In Proceedings - International Conference on Distributed Computing Systems*. Volume 2016-August, pp767–768.
- Costa, P. Grey Box City: Building cybernetic urban systems for smarter simulations. Data - SMART CITIES - Volume 1 - eCAADe 37 / SIGraDi 23. [papers.cumincad.org]
- Grammenos, F., Tasker-Brown, J. (2002). Residential Street Pattern Design. Socio-economic Series, 75. [irbnet.de] [web.archive.org]
- Hopman, R., Scott, E., Spencer, C., Weissinger, K., Strayer, D. (2017). *Measuring Cognition in Nature - Neural Effects from Prolonged Exposure to Nature*. [researchgate.net]
- Kallinikos, J. (2007). *The consequences of information*. Edward Elgar Publishing, Cheltenham.
- Kallinikos, J. (2011). *Governing through technology: information artefacts and social practice*. Palgrave Macmillan, Basingstoke.
- Kuzina, A. (2019). Conception of the operational information model of smart city control system. E3S Web of Conferences 97, 01024. DOI: 10.1051/e3sconf/20199701024 [e3s-conferences.org]
- Marshall, S. (2005). *Urban Pattern Specification*. Institute of Community Studies, London. [citeseerx.ist.psu.edu]
- Marvin, S., Luque-Ayala, A. (2017). *Urban Operating Systems: Diagramming the City*. International Journal Of Urban And Regional Research. DOI:10.1111/1468-2427.12479 [onlinelibrary.wiley.com]
- Mouhoubi, N., Boudemagh, S.S. (2015). The "Project" Approach in Urban: A Response to Uncertainty. World Academy of Science, Engineering and Technology International Journal of Humanities and Social Sciences, 9(4). [publications.waset.org]
- Ni, P., Kamiya, M., Ding, R. (2017). *Cities Network Along the Silk Road: The Global Urban Competitiveness Report 2017*. Springer.
- Oppezzo, M., Schwartz, D.L. (2014). *Give Your Ideas Some Legs: The Positive Effect of Walking on Creative Thinking*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(4), pp 1142–1152. DOI:10.1037/a0036577 [apa.org]
- Pemberton, R. (1854). *The Happy Colony*. London. From: Rowe, C. (1976). The Mathematics of the Ideal Villa and Other Essays. The MIT Press, Cambridge, Massachusetts and London, England. p81. ISBN 0-262-18077-4 [flickr.com]
- Ramaprasad, A., Sánchez-Ortiz, A., & Syn, T. (2017). *A Unified Definition of a Smart City*. Electronic Government, 13–24. doi:10.1007/978-3-319-64677-0_2
- Ramaprasad, A., Sánchez-Ortiz, A., & Syn, T. (2017). *A Ontological Review of Smart City Research*. Twenty-third Americas Conference on Information Systems, Boston. [researchgate.net]
- Ronkko, E., Herneoja, A., Oikarinen, E. (2018). *Cybernetics and the 4D Smart City: Smartness as Awareness*. Challenges, 9(1). DOI: 10.3390/challe9010021 [mdpi.com]
- Santana, E.F.Z., Chaves, A.P., Gerosa, M.A., Kon, F., Milojicic, D.S. (2017). *Software Platforms for Smart Cities: Concepts, Requirements, Challenges, and a Unified Reference Architecture*. arXiv:1609.08089 [arxiv.org]
- Stepandic, J., Wognum, N., Verhagen, N.J.C. (Eds.) (2019). *Systems engineering in research and industrial practice*. Springer. p236.
- Stepanek, P., Mouzhi, G. (2019). *Validation and Extension of the Smart City Ontology*. Proceedings of the 20th International Conference on Enterprise Information Systems (ICEIS 2018). DOI:

10.5220/0006818304060413 [scitepress.org]
- Townsend, A.M. (2000). *Life in the real-time city: mobile telephones and urban metabolism*. Journal of Urban Technology 7.2, 85–104.
- Townsend, A.M. (2015). *Cities of data: examining the new urban science*. Public Culture 27.2 76, 201–12.
- Trappey, A.J.C., Trappey, C.V., Govindarajan, U.H., Sun, J.J. Chuang, A.C. (2016). *A Review of Technology Standards and Patent Portfolios for Enabling Cyber-Physical Systems (CPS)* in Advanced Manufacturing. IEEE Access (Volume: 4). pp7356-7382. DOI: 10.1109/ACCESS.2016.2619360. [ieeexplore.ieee.org] [docplayer.net]
- Uribe-Perez, N. and Pous, C. (2017). *A novel communication system approach for a smart city based on the human nervous system*. Future Generation Computer Systems, 76:314–328.
- Voutsina, K., J. Kallinikos and C. Sorensen (2007). *Codification and transferability of IT knowledge*. In Proceedings of the 15th European Conference on Information Systems, University of St. Gallen, St. Gallen

Book references

- Alter, A. (2014). *Drunk Tank Pink: And Other Unexpected Forces That Shape How We Think, Feel, and Behave*. Penguin Books.
- Amstutz, A.E. (1968). *City management––a problem in systems analysis*. Sloan School of Management, MIT, Cambridge, MA.
- Batty, M. (2013). *The new science of cities*. MIT Press, Cambridge, MA [complexcity.info, earlier article]
- Curl, J.S. (2019). *Making Dystopia*. Oxford University Press
- Clifford, M.A. (2018). *Your Guide to Forest Bathing: Experience the Healing Power of Nature*. Red Wheel. [shinrin-yoku.org]
- Easterling, K. (2014). *Extrastatecraft: the power of infrastructure space*. Verso, London.
- Espino, N.A. (2017). *Building the Inclusive City*. Routledge.
- Forrester, J.D. (1961). *Industrial dynamics*. Pegasus Communications, Waltham, MA.
- Forrester, J.D. (1969). *Urban dynamics*. MIT Press, Cambridge, MA
- Fresco, J., Meadows, R. (2007). *Designing the future*. Osmora Publishing. [files.thevenusproject.com]
- Glanville, R. (2007). *Try again. Fail again. Fail better: The cybernetics in design and the design in cybernetics*. Kybernetes, 36, 1173–1206. [asc-cybernetics.org]
- Goodall, B. (1987). *The Penguin Dictionary of Human Geography*. London: Penguin.
- Howard, E. (1898). *Garden Cities of To-Morrow*. Swan Sonnenschein & Co. [en.wikisource.org]
- Lasker, G.E. (1981). *Systems approaches in computer science and mathematics, Proceedings of the International Congress on Applied Systems Research and Cybernetics*. Pergamon Press.
- Light, J.S. (2003). *From warfare to welfare: defense intellectuals and urban problems in cold war America*. JHU Press, Baltimore, MD
- Meier, R.L. (1962). *A communications theory of urban growth*. MIT Press, Cambridge, MA.
- Ross, A.W. (1957). A*n Introduction to Cybernetics. In Principia Cybernetica Web (Principia Cybernetica, Brussels)*. Chapman & Hall Ltd. London, UK. [pespmc1.vub.ac.be]
- Webber, M. (1964). *Explorations into urban structure*. University of Pennsylvania Press, Philadelphia
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Wiley & Sons: New York, NY.

Online references

- Batty, M. *A science of cities.* Accessed: January 21, 2020. [complexcity.info]
- *Central place theory.* (2019). Wikipedia. [en.wikipedia.org]
- *Circular cities.* The Venus Project. Accessed: January 18, 2020. [thevenusproject.com]
- Delen. (2014). *The Quadralectic Theory*. Quadralectic Architecture. [quadralectics.wordpress.com]
- Delen. (2016). *The circular/radial model*. Quadralectic Architecture. [quadralectics.wordpress.com]
- Ellis, C. *History Of Cities And City Planning.* Accessed: January 17, 2020. [art.net]
- *FAQ*. The Venus Project. Accessed: January 18, 2020. [thevenusproject.com]
- Hakan. *Radial Grid*. Monovektor. Accessed: January 18, 2020. [monovektor.com]
- Levinson, D. *Grids Are for Squares: 3 Reasons to Consider Alternatives for City Design*. Smart Cities Dive. Accessed: January 20, 2020. [smartcitiesdive.com]
- Notaro, A. (2005). *Imagining the Cybernetic City: The Venus Project. Nebula*. [cdn.atria.nl]
- Pumain, D. (2004). *Central places theory*. Hypergeo. [hypergeo.eu]
- *Quotes from Robert Park - Ernest Burgess Roderick McKenzie and Louis Wirth: Concepts for Human Ecology*. Andrew Roberts. Accessed: January 20, 2020. [studymore.org.uk]
- *Smart city.* (2020). Wikipedia. [wikipedia.org]
- *Systems of cities*. (2009). The World Bank: The International Bank for Reconstruction and Development. [siteresources.worldbank.org]
- *The circle*. (2011). AMSI. [amsi.org.au]
- *The Origin of Cities – Part 1*. (2016). the HipCrime Vocab. [hipcrimevocab.com]
- *The Origin of Cities – Part 2*. (2016). the HipCrime Vocab. [hipcrimevocab.com]
- *The Origin of Cities – Part 3*. (2016). the HipCrime Vocab. [hipcrimevocab.com]
- *The Origin of Cities – Part 4*. (2017). the HipCrime Vocab. [hipcrimevocab.com]
- *The Real Reason Jaywalking Is A Crime (Adam Ruins Everything)*. (2015). CollegeHumor. [youtube.be]

TABLES

Table 1. City > Ontology: *Smart city ontology within a community-type society.*

Smart								City		
Structure		**Functions**		**Focus**		**Semiotics**		**Stakeholders**		**Outcomes**
Information Technology		Sense		Life		Data		Users		Values (e.g., sustainability, resilience, etc.)
Projects		Monitor		Technological		Information		Teams		Quality-of-life / Well-being
Teams		Process		Exploration		Knowledge				Fulfillment
Processes	[to]	Translate	+	Resources	+		[by/from/to]		[for]	Flourishing
Procedures		Communicate		Access						Flow
				Social						
				Decision						
				Lifestyle						
				Material						

Land Accounting System

Travis Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

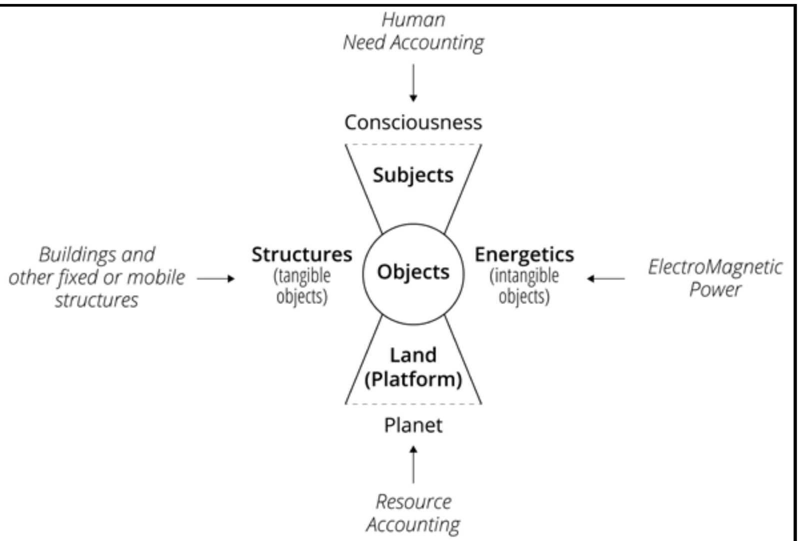
Keywords: land, land accounting, land assessment, geo accounting, site analysis, site survey

Abstract

All earth-based platforms must account for land. When engineering anything on land or ocean, the environment must be accounted for. Land is one necessary element for which to account. Land is accounted for through a land/site survey. When in use, land may be zoned and accounted for via the service/function upon its geo-coordinated position. In the market, land is considered property. In the State, land is considered jurisdiction.

Graphical Abstract

Figure 10. *Depiction of the account of material objects composed of four primary categories (structures, subjects, energetics, and land), of which land is a fundamental component. It is upon land (or, a "landed platform") that a set of useful services may be sustained.*



1 Introduction

Site analysis is an surveying and analytical process that gathers information from an environment where probable operation is to occur. Primarily, it involves the collection of data from various sources about a given spatial area where placement may occur. Data may be collected from all primary mediums in the placement area (i.e., land, liquid, and atmosphere). Site assessment is necessary in every design processes where placement in space is a consideration. It involves the gathering of data from a site for use in site selection and engineered construction.

Terminology includes, but is not limited to:

- Topography is a detailed map of the surface features of land [contours]. Topology is the study of place, from topo-, combination form of Greek topos "place" + -logy "study of".
- Site analysis
- Natural systems
- Human systems

PRIMARY: *For anything which is to be built, its design must account for its placement.*

In concern to the placement of buildings and other technologies on a specific area of land, then a site inquiry assesses the geomorphology and climate of the given environment [in the context of that which is to be selected or built].

For example, you would want to know whether the placement of a garden is on top of a former dumping ground. Or, you would want to know the surface shape upon which an object is to be placed -- is the surface round or flat, and how smooth?

A site analyses can provide data to analyse the difference (compare and contrast) between possible placement site locations.

A site survey provides information that may be useful in decisions involving:

- Site selection;
- The design of the object to be placed; and
- The re-design of the spatial area into which the object is to be placed.

The typical phases/generic steps in site analysis are program investigation, site investigation and analysis, site evaluation, and report development.

1. **Program investigation:** The building program is investigated with respect to the selected or optional building footprints; area required for parking, circulation, open space, and other program elements; and any special constraints or requirements such as security, easements, preserving natural habitat, wetlands, and the like.
2. **Site inventory and analysis:** The physical, cultural, and regulatory characteristics of the site are initially explored. The site evaluation checklist identifies factors that may be considered. Some of these factors can be assessed by collecting and analyzing information; others are best addressed by walking the site and traversing its environs. A preliminary assessment of whether a location and site have the potential to accommodate the building program is made. Priority issues—those (such as environmental contamination) that may preempt further investigation—are identified. A site analysis plan is developed. When this has been approved by the client, consultants may be hired to further explore issues that require analysis beyond the capabilities of the core project team.
3. **Site evaluation:** At this point, thorough assessments are conducted when necessary to develop the site analysis plan. These may include physical testing of aspects of the site, its improvements, and adjoining properties.
4. **Report development:** The site analysis report normally includes property maps, geotechnical maps and findings, site analysis recommendations, and a clear statement of the impact of the findings and recommendations on the proposed building program.

NOTE: *Regulatory approvals normally required during or immediately following the site analysis phase include zoning, environmental impact, and utilities & transportation.*

1.11.1 Geographical information system

Geographic information system (GIS) refers to the process collecting and mapping data about a spatial-temporal location. A GIS process records spatio-temporal (space-time) location as the index variable for all information.

In other words, just as a relational database containing text or numbers can relate many different tables using common index variables, GIS can relate otherwise unrelated information by using location and time as the index variable. The contextualizing factor is the location and/or extent in space-time. Any variable that can be located spatially, and increasingly also temporally, can be referenced using a GIS.

Locations or extents in space–time may be recorded as dates/times of occurrence, and x, y, and z coordinates representing, longitude, latitude, and elevation, respectively. Earth-based spatial–temporal location and extent references should, ideally, be relatable to one another and ultimately to a "real" physical location or extent.

Some common GIS data about any given location on

Earth might include topology with contour lines and elevations, soil and land mass composition, land use, wildlife, and even political districts. Any technology that can sense the real world can be used to collect GIS data. Some of the most common technologies referred to in the discussion of a GIS are the global positioning system, atomic clocks, remote sensing and imaging systems. Some commonly included models in a GIS are hydrological modeling, cartographic modeling.

2 Land survey assessment

3 Site survey assessment

This is a checklist of the factors that may be involved in the gathering of data about a "site". A "site" is a spatial location of an area where something which is being designed and constructed may likely, or will, be placed). In our controlled habitat information is continuously collected, because the designed construction of our habitat service system is emergent.

> **NOTE:** *The site survey assessment as it presently exists here is oriented toward market-State conditions and does not represent a clean site survey as would be done in an environment without the market-State.*

3.1 Physical factors

3.1.1 Climate

1. Prevailing winds
   A. Direction
   B. Maximum, minimum, and average velocities
   C. Special forces (e.g., tornadoes, hurricanes)
2. Solar orientation
   A. Sun angles
   B. Days of sunlight
   C. Cloud cover
   D. Shading of (or from) adjacent structures, natural features, vegetation
3. Temperature
   A. Ranges of variation
   B. Maximums and minimums
4. Humidity
   A. Ranges of variation
   B. Maximums and minimums
   C. Precipitation
   D. Peak period totals
   E. Annual and seasonal totals

3.1.2 Topographic factors

3.1.2.1 Land topography

5. Topographic maps and aerial images
   A. Contours and spot evaluations
   B. Slopes; percentage, aspect, orientation
   C. Escarpments
   D. Erosion channels
   E. Extent, location, and general configuration of rocks, ledges, outcrops, ridges, drainage lines, and other unique features
   F. Visual characteristics
   G. Potential problem areas during construction: situation, erosion, precipitation, etc.
6. Analysis of physical features, including major focal and vantage points and their relationship within, into, and out from the site.
7. Existing access and circulation
   A. Human locomotion
   B. Vehicle locomotion
8. Vegetation
9. Existing water bodies
   A. Location, size, depth, direction of flow
   B. Water quality: clean, polluted, anaerobic conditions, etc.
   C. Use: seasonal, year-round
   D. Wetlands: ecological features
   E. Variations: expected water levels, tides, wave action
   F. Coastal features
10. Drainage canals: rivers, streams, marshes, lakes, ponds, etc.
    A. Natural and built
    B. Alignments and gradients
    C. Patterns and direction
11. Existing waterway easements
    A. Surface
    B. Subsurface
12. Surface drainage
    A. Patterns on and off the site (location of streams and washes)
    B. Proximity to floodplains: a) maximum flood levels; b) frequent flood areas
    C. Local watershed areas, amount of runoff collected, and location of outfalls
    D. Swampy and concave areas of land without positive drainage and other obstacles that may interrupt or obstruct natural surface drainage
    E. Potential areas for impoundments, detention/ retention ponds
13. Unique site features

3.1.2.2 Geotechnical/soils

14. Basic surface soil type: sand, clay, silt, rock, shale, gravel, loam, limestone, etc.
15. Rock and soil type: character/formation and origin
    A. Geologic formation process and parent material
    B. Inclination
    C. Bearing capacity
16. Bedrock
    A. Depth to bedrock
    B. Bedrock classification
17. Seismic conditions
18. Environmental hazards

3.1.2.3 Utilities

19. Portable water
20. Electricity

21. Gas
22. Telephone
23. Cable television
24. Sanitary sewer service
25. Storm drainage (surface, subsurface)
26. Fire protection

3.1.2.4 Immediate surroundings

27. Neighbourhood structures: buildings, cell towers, airports, flight paths, satellite dishes, etc.
28. Shading and solar access
29. Noise from streets, emergency services, aircraft, etc.
30. Odors
31. Views and vistas

3.1.2.5 General services

32. Fire and police protection
33. Trash/refuse removal services
34. Snow removal, including on-site storage

3.2 Prior land use factors

3.2.2.6 Site history

35. Former site uses
    A. Hazardous dumping
    B. Landfill
    C. Old foundations
    D. Archaeological grounds
36. History of existing structures
    A. Historic worth
    B. Affiliations
    C. Outline
    D. Location
    E. Floor elevations
    F. Type
    G. Condition
    H. Use or service

3.2.2.7 Land use, ownership, and control

37. Analysis of legal property, including limits of property, easement, rights of way, and north indication
38. Present zoning of site and adjacent property
39. Adjacent (surrounding) land uses
    A. Present
    B. Projected
    C. Probable effects on the development of this site
40. Type of land ownership
41. Function and pattern of land use: public domain, farm type, grazing, urbanized
    A. Present
    B. Former
42. Location, type, and size of pertinent community services
    A. Schools and churches
    B. Shopping centers
    C. Parks
    D. Municipal services
    E. Recreational facilities
    F. Banks
    G. Food services
    H. Health services
    I. Access to highways, public transportation

3.2.2.8 Economic value

43. Political jurisdictions and land costs
44. Accepted “territories”
45. Future potential
46. Size of surrounding lots and approximate price ranges

3.3 Regulatory factors

3.3.2.9 Zoning codes

47. Permitted uses
    A. By variance
    B. By special use permits
    C. Accessory structures
48. Minimum site area requirements
49. Building height limits
50. Yard (setback) requirements
51. Lot coverage
    A. Floor area ration (FAR)
    B. Percentage of coverage
    C. Open space requirements
52. Off-street parking requirements
53. Landscaping requirements
54. Sign requirements

3.3.2.10 Subdivision, site plan review, and other local requirements

55. Lot requirements
    A. Size
    B. Configuration
    C. Setbacks and coverage
56. Street requirements
    A. Widths
    B. Geometry: grades, curves
    C. Curbs and curb cuts
    D. Road construction standards
    E. Placement of utilities
    F. Dead-end streets
    G. Intersection geometry
    H. Sidewalks
    I. Names

57. Drainage requirements
    A. Removal of spring and surface water
    B. Stream courses
    C. Land subject to flooding
    D. Detention/retention ponds
58. Parks
    A. Open space requirements
    B. Park and playground requirements
    C. Screening from adjacent uses

3.3.2.11 Environmental regulations

59. Water, sewer, recycling, solid waste disposal
60. Clean air requirements
61. Soil conservation
62. Protected areas, wetlands, floodplains, coastal zones, wild and scenic areas
63. Fish and wildlife protection
64. Protection of archaeological resources

3.3.2.12 Other codes and requirements

65. Historic preservation and landmarks
66. Architectural (design) controls
67. Special districts
68. Miscellaneous, e.g., mobile homes, billboards, noise
69. Site-related items in building codes
    A. Building separation
    B. Parking and access for persons with disabilities
    C. Service and emergency vehicle access and parking

TABLES

Table 2. Land Accounting*: Hierarchical classification of geomorphological features (time and space scales are approximate).*

Hierarchical Classification Of Geomorphological Features			
Typical units		**Spatial scale km^2**	**Time Scale Years**
Continents		10^7	10^8-10^9
Physiographic provinces, mountain ranges		10^6	10^8
Medium and small scale units, domes, volcanoes, troughs		10^2-10^4	10^7-10^8
Erosional/depositional units:			
	Large scale, large valleys, deltas, beaches	10-10^2	10^6
	Medium scale, floodplains, alluvial fans, cirques, moraines	10^{-1}-10	10^5-40^6
	Small scale, offshore bars, sand dunes, terraces	10^{-2}	10^4-10^5
Geomorphic process units:			
	Large scale, hillslopes, channel reaches, small drainage basins	10^{-4}	10^3
	Medium scale, slope facets, pools, riffles	10^{-6}	10^2
	Small scale, sand ripples, sand grains, striations	10^{-8}	

Table 3. Land Accounting*: The earth's water/hydrological cycle processes.*

The earth's water cycle processes	
Process	**Definition**
Precipitation	Condensed water vapor that falls to the Earth's surface . Most precipitation occurs as rain, but also includes snow, hail, fog drip, graupel, and sleet.
Canopy interception	The precipitation that is intercepted by plant foliage, eventually evaporates back to the atmosphere rather than falling to the ground.
Snowmelt	The runoff produced by melting snow.
Runoff	The variety of ways by which water moves across the land. This includes both surface runoff and channel runoff. As it flows, the water may seep into the ground, evaporate into the air, become stored in lakes or reservoirs, or be extracted for agricultural or other human uses.
Infiltration	The flow of water from the ground surface into the ground. Once infiltrated, the water becomes soil moisture or groundwater. A recent global study using water stable isotopes, however, shows that not all soil moisture is equally available for groundwater recharge or for plant transpiration.
Subsurface flow	The flow of water underground, in the vadose zone and aquifers. Subsurface water may return to the surface (e.g. as a spring or by being pumped) or eventually seep into the oceans. Water returns to the land surface at lower elevation than where it infiltrated, under the force of gravity or gravity induced pressures. Groundwater tends to move slowly, and is replenished slowly, so it can remain in aquifers for thousands of years.
Evaporation	The transformation of water from liquid to gas phases as it moves from the ground or bodies of water into the overlying atmosphere. The source of energy for evaporation is primarily solar radiation. Evaporation often implicitly includes transpiration from plants, though together they are specifically referred to as evapotranspiration.
Sublimation	The state change directly from solid water (snow or ice) to water vapor.
Deposition	This refers to changing of water vapor directly to ice.
Advection	The movement of water — in solid, liquid, or vapor states — through the atmosphere. Without advection, water that evaporated over the oceans could not precipitate over land.
Condensation	The transformation of water vapor to liquid water droplets in the air, creating clouds and fog.
Transpiration	The release of water vapor from plants and soil into the air. Water vapor is a gas that cannot be seen.
Percolation	Water flows vertically through the soil and rocks under the influence of gravity.
Plate tectonics	Water enters the mantle via subduction of oceanic crust.

Resource Accounting System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: resource, resource accounting, object accounting, resource survey

Abstract

All earth-based platforms must account for resources. When engineering anything, its object composition, and the composition of the environment where objects are distances from one another, must be accounted for. Resources are one necessary element for which to account when planning project and operationalizing a product. Resources are accounted for through a global resource survey. In the market, resources are property. Spatial resources (true resources) are objects (i.e., made of matter/shape). Informational "resources" (digital resources) are data (i.e., made of bits). Human "resources" (contributors) are individuals (i.e., made of consciousness). Consciousness uses spatial and informational resources in order to sustain its embodiment and to develop itself.

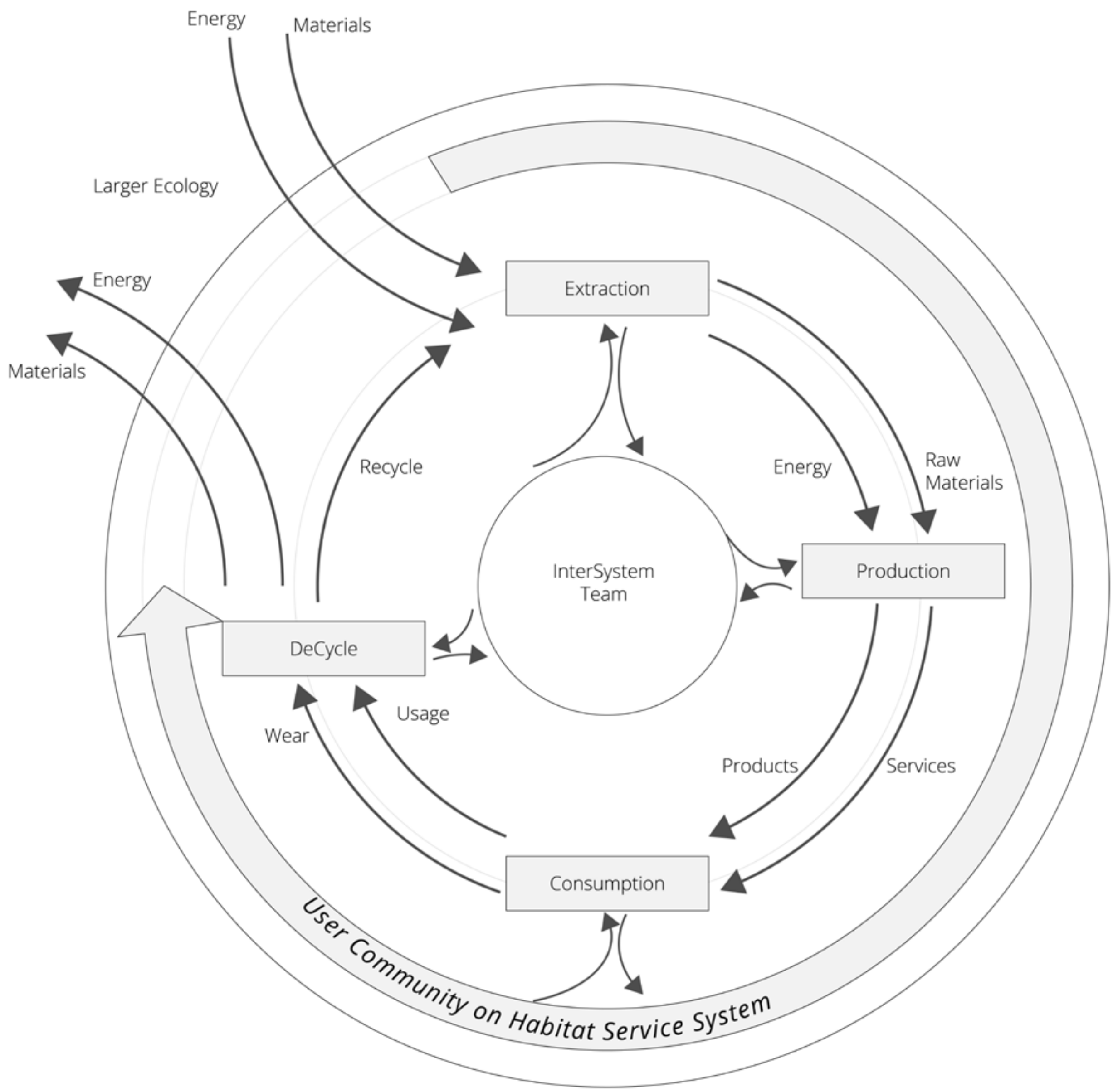
Graphical Abstract

Figure 11 on page 52

1 Resource surveying

A.k.a., Global resource survey.

Figure 11. *The engineering of a real-world material cycle of resources for human "consumption" through team contribution.*



Access Accounting System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: access, access accounting, access surveying

Abstract

All societal-based platforms must account for access. When producing anything, access to that thing must be accounted. Access is necessary and two dimensional concept. Firstly, there is access to a team or working group through a contribution-based structure, and then, there is access to goods and service (without force of trade). Access can be accounted for many types of surveys including demand surveys, contribution surveys, etc. In the market, access is considered through the cost of a sale. In the State, access is acquired through authority. Humans require access to objects and information, which are composed into services. In a market, access is controlled by price, and the concept itself is mixed with "rights" (given by authority) and "property" (purchased in the market). In a community-type society, access refers to demands and other issues for service that are accessible to users. Ultimately, the goal is to have access to that which optimally meets user requirements (human needs) given that which is available at the time of access. In a community-type society, access centers and integrated transportation systems distribute products. Services are integrated, often modularly, into the infrastructure of the environment in order to optimize efficiency and produce a higher quality experience of access [to services] by a user. With sufficient technical knowledge and ability it is possible to apply automation technologies to increase the efficiency by which access occurs. Automation technologies can free individuals for access to opportunities they might otherwise not have had. Automation technologies can also make access to services, such as medical and informational more safe, reliable, and faster.

Graphical Abstract

Figure 12 on page 55

1 Access centers

A.k.a., Libraries, repositories, etc.

A library of accessible items are available at access centers.

2 Automated storage and retrieval system

No content here yet *(sometimes shown and sometimes not; not yet complete, but category is known)*

3 Access location designations

There are three types of material access designation:

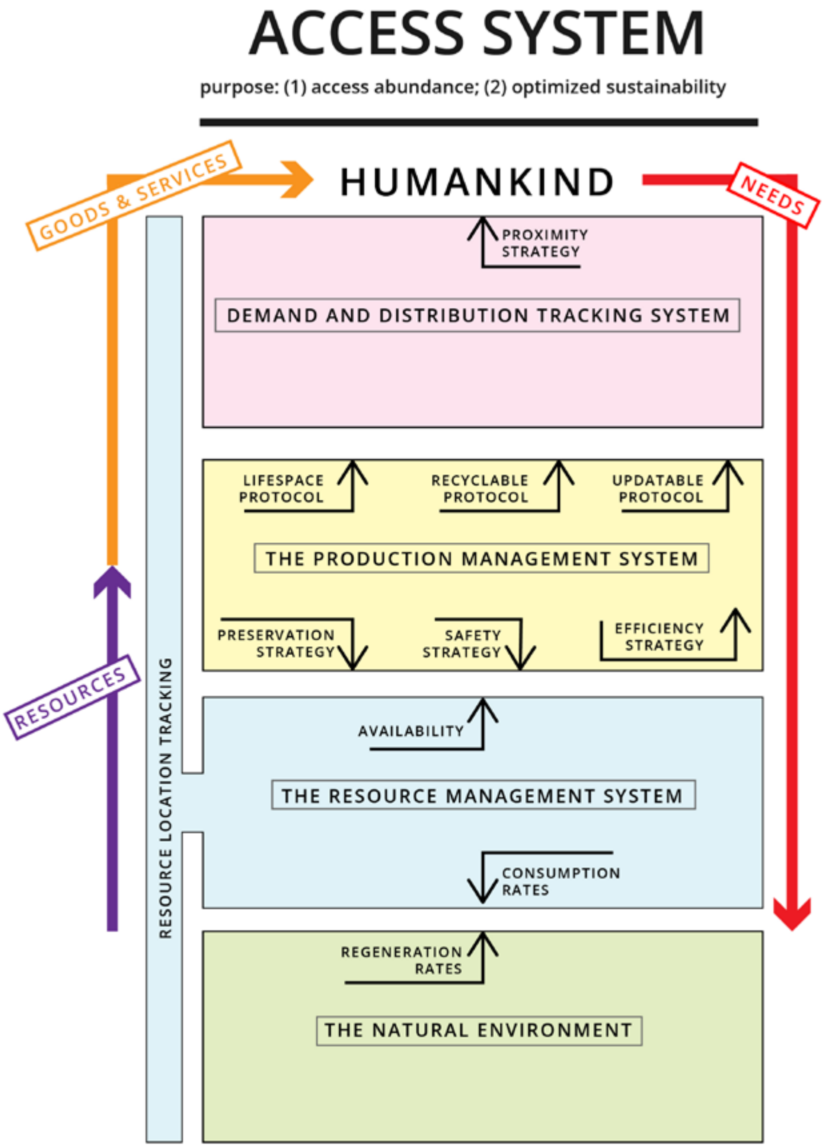
1. Personal access designations
2. Common access designations
3. Systems access designations

Location specific [visitation] attributes include, but are not limited to:

1. Frequency [of visit to location]
2. Duration [of visit to location]
3. Recency [of visit to location]

These elements should be accounted for in the design and selection of zones, circulars, and other material access locations.

Figure 12. *Resource tracking within an access system.*



Technology Accounting System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

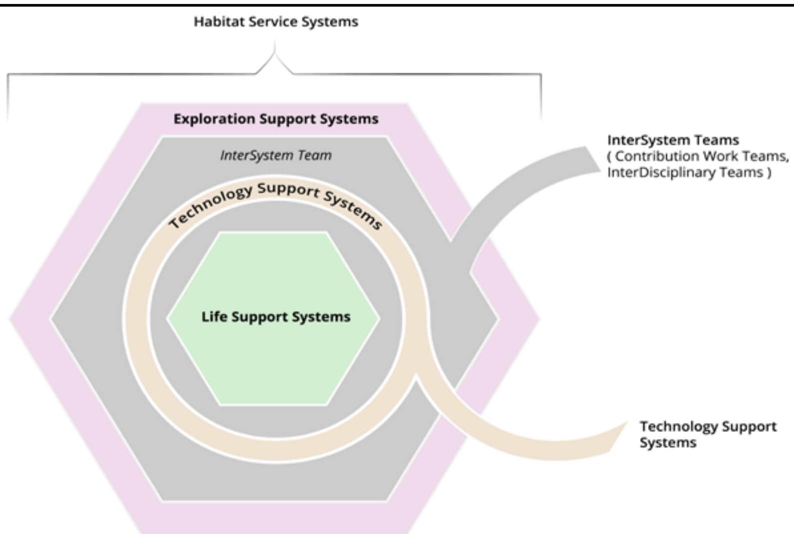
Keywords: technology, technology accounting, technology readiness, technology development, technology operation,

Abstract

All technical-based platforms must account for technology. Society, as a socio-technical system, must account for technology. Technology is an enabling element in society; it enables doing more with less. Technology can be accounted for through technology surveys. In the market, technology may be is considered defensible property. In a community-type society, technology is a tool for well-being. Technology is accounted for in order to optimize societal configuration. Technologies can be compared between. A solution within the decision system can compare the technical requirements and consequences of different technologies. The orientational value system within the decision system will compare analyze the technologies in relation to formalized desirable conditions. Some technologies are less well understood and developed than others. Technology can be identified by the manner in which the technology is actually used by individuals. Technology can be accounted for by the work done by scientists and engineers (as developers). The diffusion of technology through cities in a given society can be accounted for, by at least, resource availability, knowledge availability, and local [cultural] values.

Graphical Abstract

Figure 13. *The integration of technology into a habitat service platform for human human fulfillment, involving (at least) life and exploration support.*



1 What is technology

> **NOTE:** *Community articulates reasons for using technology, while applying technology systematically for human and ecological benefit. In early 21st century society technology is selected and applied ad hoc and for profit.*

The word technology comes from the Greek word "tekhnologia" (or, "tekne"), which means: art, skill, craft in work; method, system, an art, a system or method of making or doing. "Craft" originally meant "weaving" or "fabricating". Today, technology may be defined as (i.e., technology is):

1. An expression of meaning through values that encodes functionality into a specific configuration of matter that processes information as energy. Technology is an expression of meaning that expresses values and encodes functionality in a specific, materially usable system.
2. The intentional application of knowledge. If technology is applied knowledge, then habitat service system represent the applied encoding of technology into a habitat.
3. The applied result of the logical ordering of factual technical relationships; technology is applied knowledge.
4. The practical application of knowledge into a construction for a function. It refers to a technological process, method, or technique such as machinery, equipment or software needed for a service or process to achieve its purpose. Even a practice or process can be a technology, although in common parlance, people have a tendency to think of technology solely as something which is material.
5. Some thing with a function or utility to it. Technology is simply the application of [accurate] knowledge toward the extension of function.
6. A systematic creation that allows an organism to do a task more efficiently than using its body alone.
7. An engineered creation.
8. The result of effectively re-structuring an environment through intention.
9. A material object with measurable dimensions and material attributes.
10. A procedural model (or procedural object) with measurable affects on information and conceptual attributes.

Technology is [systematically] built on prior verified and understood patterns of relationship in a commonly objective, existent reality. Technical processes underlay every designed structure in the real world. At a societal level, technology is that which is tested and works.

All technological systems are:

1. Purposeful.
2. Built on/from technical principles.
3. Subject to off-normal events, including accidents and faults, due to both component malfunctions and unforeseen/foreseen external influences. Therein, some technology requires additional technology to minimize the likelihood and impact of accidents.

The primary functions of technology for a population include, but are not limited to:

1. Providing life support service.
2. Providing exploratory support services.
3. Providing for production (and materials re-cycling), communication, transportation, and information processing.
4. Making more efficient use of time.

The authentic use of science and technology accords many benefits, including but not limited to:

1. Improving the quality of lives by automating banal labor tasks.
2. Creating more intuitive, natural, and active learning and information systems.
3. Improving the quality of decisions and problem solving.
4. Improving the effectiveness and efficiency with which goods and services are distributed.
5. Supporting in the creation of an abundance of all goods and services. Among the most important benefit of science and technology in our community is that derived from technological automation. It is the effort to free the individual from banal labor contributions so that they may pursue their own interests, improving both themselves and the community.

> **NOTE:** *The human organism is a technological construction; a technology for the animated expression of consciousness in material form. Wherein, consciousness has a technical and intuitive relationship to the form it inhabits.*

1.1 Technology sub-classification

> **INSIGHT:** *Technology is not a panacea. However, it can be extremely useful in solving many kinds of problems.*

Technologies can be sub-classified according to:

1. **Structure** - The structure is the components of the design object and their relationships. A system

is a structured form of organization. The structure (a.k.a., architecture) of the system designed to transform information for a purpose.

2. **Process** - The occurrence of an operational transformation (or event). A process produces a behavior [for a specific function]. The behavioral process(es) represents the attributes (or "qualities") that can be derived from the designed object's structure. A system is a form of organization that includes at least one process.
3. **Function** - The objective [purpose or goal] for the transformation within and overall existence of the system. A system is a functional form of organization.
4. **Materials** - The specific material(s) that compose the technical system (i.e., material composition).

1.2 Tools

A tool is a device that is necessary to, or aids in the performance of an operation. Tools are the manifestation and extension of consciousness. It is possible to explain how work (Read: an operation) is done through 'tools'. In the production and use of a tool there is the opportunity to apply a strategy that orients the use of a tool in a particular direction. Someone can hammer a nail into wood, but hammering continues past a particular point, then a divot may be created, and then, a hole. All tools have a contextual and orientational use. A tool's value is "put there" by the human or system that uses it.

> **Tool**
> *Noun*
> *1. A device or implement, especially one held in the hand, used to carry out a particular function.*

Tools are a structured part of a society's environment that supports sense-making, enables engineering, and facilitates servicing. There are both conceptual (informational) tools as well as physical (object) tools. A material tool is an object used to extend the ability of an individual to modify features of the surrounding material environment. A tool is the most basic type of ability extender and is comprised of a resources. A conceptual tool is a concept [model] used to extend the ability of an individual to modify features of the surrounding informational environment. Informational tools include models, programs, and algorithms.

> **CLARIFICATION:** *The terms tool and technology are often used synonymous, in other cases, a tool is a specific application or instance of a technology. In some cases, tools are physical and technologies are information.*

1.3 Technology and morality

> **INSIGHT:** *When you invent a technology you also invent the accident and/or misuse of that technology.*

Technology is amoral. What humanity chooses to do (or, not to do) with technology is a moral choice; the difference humanity creates in its environment is a moral (or "ethical") decision, that can be resolved via a direction, orientation, and approach to its decisioning. It is what technologies we create and what we do with technology that makes it moral (fulfilling) or immoral (unfulfilling).

When we perceives technologies taking us in "dangerous" directions, consider that maybe it is really our way of life, and our lifestyle, that is taking us in a dangerous direction. We are just using technologies in ways that we couldn't before (because technology is allowing us to do more of what we are doing already), and that is where there is danger.

Engineered creations will take on the biases and standards, the intentions, of the socio-economic system in which they have been designed and will be utilized. Technologies created and applied in a capitalist system will have a capitalist bias. Alternatively, technologies created and applied in community will likely maintain standards that orient all of humanity toward greater fulfillment and clarity of perception.

Human fulfillment is more important than technological progress and innovation. Societies that prioritize technological advancement over human fulfillment are likely to forget that the situations in which innovation and technological advancement are likely to align with fulfillment are the situations where creators and users can freely and visibly decide upon which innovations to encode (i.e., "incorporate") into their lives. Some societies put too much emphasis on innovation as a goal (or "economic growth" and "entrepreneurship"), which is to take for granted that innovation is always good for human well-being and fulfillment. In the market, innovations and products serve the (for profit and power) interests of market entities - they have a market/capitalist/State bias. Conceptually speaking, technology is neutral. For example, with electricity "you" can kill someone or make dinner for someone. However, specific technologies can be evaluated as setting up conditions that orient more greatly toward or away from fulfillment. For example, technologies that malfunction easily are likely to setup suboptimal resource usages, which has consequence for societal orientation as a whole. A certain kind of technology and/or its implementation can be seen to lead to certain kinds of effects. A given kind of technology can establish conditions that are not morally neutral. In a community-type society, technology is developed and applied through community-based value standards.

One might put oneself in the position of technology itself, and then ask oneself, “What would I do if I was this technology?” For example, "What would I do if I were a coffee maker, a bridge, a rifle, or a nuclear bomb?" We can imagine what these technologies would do if they “wanted” to be applied.

It must be noted here that material technology is just a piece of the modern sustainability puzzle. Solving the systemic challenges facing our global community requires context and an accurate value orientation in addition to sustainable conceptual and technological solutions. Technology is not sufficient to fix our problems; we need a moral organizational architecture. Technology without morality and intrinsic motivation is likely to promote apathetic idleness.

2 The technology development matrix

A.k.a., Technology readiness matrix, technology readiness level, etc.

The technology development matrix (TDM) is an interactive visual matrix designed to track and facilitate the development of those technologies that are necessary for human life to thrive in community. The TDM is a checklist and meta-information source of what technologies, systems and capabilities are available to build a critical path to human fulfillment. The availability of this information allows anyone to get involved in tackling the difficulties and challenges associated with technological development. The TDM could be used as a “punch list” for building the material construction of a city in community.

By accessing the matrix anyone can see who is working on which technology. If nobody is on a technology of interest to you, then you drop in your abstract, your concept, and contact information. The matrix is a view into the technological possibilities and capabilities of the community.

Generally speaking, technology development has three characteristics:

- It is the process of developing and demonstrating new or unproven technology.
- It is the application of existing technology to new or different uses, or
- It is the combination of existing and proven technology to achieve a specific goal.

The TDM also identifies the readiness levels of each technology.

Readiness refers to time. Specifically it means ready for operations at the present time. Level refers to the level of maturity of equipment. Equipment that is already being used for the same function in the same environment has a higher level of maturity than equipment that is still being developed. The levels are a nine-point scale based on a qualitative assessment of maturity.

Technology Readiness Level (TRL) is an index to measure the development and usability of an evolving technology. It measures how ready equipment is for use now in an operating service. A technology’s “readiness level” (TRL) refers to its phase of existence. The primary purpose of using technology readiness levels (TRLs) is to inform the resolution of a decision space concerning the development and transitioning of technology. The TRL index scale goes from 1-9.

- TRL 1-3 = red, a theoretical concept
- TRL 3-6 = yellow, is being tested in the lab
- TRL 6-9 = green, is being applied

Almost all the TRL scale developers and users in various perceive TRL 6 to be a major transition from research and experiment to real life implementation and operation. At TRL 6, a representative model, prototype or system, which would go well beyond an ad hoc discrete component level breadboard, must be tested in a relevant environment. If the only relevant environment to show progress is the operational environment, then the validation must be demonstrated in operational environment. At TRL 6, several (or many) new technologies will typically be integrated into the demonstration so a working, sub-scale (but scaleable) model of the system should be successfully demonstrated.

Note, however, that the idea of a "technology readiness level" does not apply if the objective of a project is to research scientific principles.

2.1 Technology readiness levels

> **INSIGHT:** *Capitalism hasn't given "you" any technology. Instead, capitalism gives "you" the need to buy and sell technology. In capitalism, there is willful withholding of technology and efficiency to remain competitive in the market.*

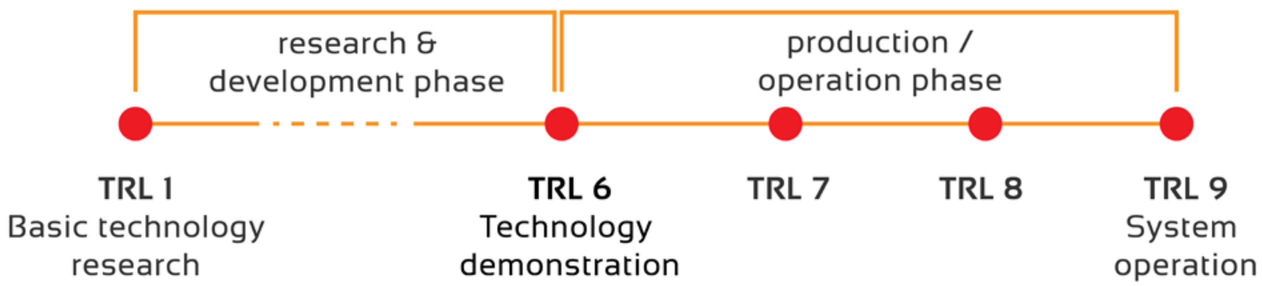
The following are reference points for reading and usage of the technology readiness level index:

1. **A TRL number is obtained once the description in the diagram has been achieved.** For example, when a technology successfully achieves TRL 5, it does not move to TRL 6. Therefore, reporting TRL 6 should be conclusively done with TRL 6 activities and validation.
2. **If a technology consists of various sub-technologies, its TRL number is the lowest of all.** A technology may depend on a number of technologies or sub-systems with their own TRLs. Then, the ultimate technology is assigned with the lowest TRL number among them.
3. **When an element of a technology is altered, its previous TRL number becomes invalid.** When one replaces, eliminates, or adds a major component or part even in a TRL-9 technology, everything starts all over again from the appropriate TRL usually between 1-4.
4. **When the primary use of a technology changes, its previous TRL number becomes invalid.** If you try to integrate (launch) a technology (product) into a different system (market), you cannot claim its previous TRL number any more. You should work through TRL validations again.
5. **If a technology spends too much time at a given TRL, its TRL number becomes invalid.** As time goes by, even a TRL 9 technology requires re-confirmation due to the probable changes in the conditions (i.e. know-how, climactic environment) that its previous TRL number is based on.
6. **Activities and progress through TRLs are not time-boxed.** Some technologies may evolve faster than others. Or, a particular technology may pass some levels in weeks but the others in years.
7. **TRL activities and validation criteria are subjected to change over time.** You cannot precisely specify TRL 8 requirements for a project while you are at TRL 2 stage and keep them the same along the way. Inspection and adaptation are needed.

Important terms to know when using the technology readiness level index include:

1. **Prototype:** A physical or virtual model used to evaluate the technical or manufacturing feasibility or utility of a particular technology or process, concept, end item, or system.
2. **Model:** A functional form of a system, generally reduced in scale, near or at operational specification. Models can be sufficiently hardened to allow demonstration of the technical and operational capabilities required of the final system.
3. **Demonstration/Pilot:** Actions aiming to validate the technical and ecological viability of a new or improved technology, product, process, service or solution in an operational (or near to operational) environment.
4. **Critical technology element:** A new or novel

Figure 14. *Technology readiness level production/operation phase.*



component that a technology or system depends on to achieve successful development or to successfully meet a system operational threshold requirement.

5. **Relevant environment:** Testing environment that simulates the key aspects of the operational environment.
6. **Operational environment:** Environment that addresses all of the operational requirements and specifications required of the final system, including platform/packaging.

2.2 Technology obsolescence levels

The general technology obsolescence levels are:

1. Obsolescence not an issue.
2. Technology is the state-of-the practice; emerging technology could integrate or replace.
3. Technology is outdated and use should be avoided in new systems; spare parts supply is scarce.

In industry, technology primarily becomes obsolete when another technology becomes more cost effective for the functionality provided -- emerging technologies compete with established technologies for market share over a particular [service] functionality. In community, new discoveries and engineered systems render old technologies obsolete. In a market, along with technology readiness there is also something called a “market readiness” level (MRL) - peoples desire and resources to purchase a technology (i.e., the readiness of people to consume the technology).

1. Technology name.
2. Technology material composition.
3. Technology material configuration.
4. Technology operational usage requirements.
5. Technology safety requirements.
6. Video of technology in use.
7. Alternative technologies.

3 Technology phasing

The issue of phasing technology into a societal system is often called, “the infrastructure problem”.

4 Technological automation

4.1 Technological productivity

Technological productivity can be calculated as:

- Productivity = what is produced / # of hrs takes to produce.

4.2 Technological unemployment

> **INSIGHT:** *Social progress can masquerades as technological advancement. Technology can lead to greater apathy and acquiescence.*

Technological unemployment is a market-State concept that refers to the occurance of technology displacing workers by substituting their cognitive and fine motor control with machines.

Scholarly references

- *Definition of the TRLs and their criteria of assessment, ISO 16290:2013.* (2013). ISO. [spacewx.com]
- *Department of homeland security science and technology readiness level calculator (ver 1.1).* (2009). Homeland Security Institute. [ansr.org]
- *Guide to Technology Readiness Levels for the NDA Estate and its Supply Chain*. (2014). Nuclear Decommissioning Authority, 2014. [assets.publishing.service.gov.uk]
- Final Report of the NASA Technology Readiness Assessment (TRA) Study Team. (2016). U.S. NASA. [ntrs.nasa.gov]
- *Technology Readiness Assessment (TRA) Guidance.* (2011). U.S. Department of Defense. [apps.dtic.mil]
- *Technology readiness assessment guide*. (2011). U.S. Department of Energy. [lbl.gov] [directives.doe.gov]
- *Technology readiness assessment guide*. (2016). U.S. GAO. [gao.gov]
- *Technology readiness levels handbook for space applications*. (2008). European Space Agency. [artes.esa.int]

Book references

- Ford, M. (2016). *Rise of the Robots: Technology and the Threat of a Jobless Future*. Basic Book.
- Ford, M. (2009). *The Lights in the Tunnel: Automation, Accelerating Technology and the Economy of the Future*. CreateSpace Independent Publishing Platform.
- Pistono, F. (2012). *Robots Will Steal Your Job, But That's OK: how to survive the economic collapse and be happy*. CreateSpace Independent Publishing Platform.

TABLES

Table 4. Technology*: Technology readiness levels.*

Number	Technology Readiness Level	Description
1	Basic principles observed and reported	Lowest level of technology readiness. Scientific research begins to be translated into technology's basic properties.
2	Technology concept and/or application formulated	Invention begins. Once basic principles are observed, practical applications can be invented. The application is speculative and there is no proof or detailed analysis to support the assumption. Examples are still limited to paper studies.
3	Analytical and experimental critical function and/or characteristic proof of concept	Active R&D is initiated. This includes analytical studies and laboratory studies to physically validate analytical predictions of separate elements of the technology. Examples include components that are not yet integrated or representative.
4	Component and/or breadboard validation in laboratory environment	Basic technological components are integrated to establish that the pieces will work together. This is relatively "low fidelity"compared to the eventual system. Examples include integration of "ad hoc" hardware in a laboratory.
5	Component and/or breadboard validation in relevant environment	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements so that the technology can be tested in simulated environment. Examples include "high fidelity" laboratory integration of components.
6	System/subsystem model or prototype demonstration in relevant environment	Representative model or prototype system, which is well beyond the breadboard tested for level 5, is tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high fidelity laboratory environment or in a simulated operational environment.
7	System prototype demonstration in an operational environment	Prototype near or at planned operational system. Represents a major step up from level 6, requiring the demonstration of an actual system prototype in an operational environment. Examples include testing the prototype in a test bed aircraft.
8	Actual system completed and qualified through test and demonstration	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this level represents the end of true system development. Examples include developmental test and evaluation of the system in its intended weapon system to determine if it meets design specifications.
9	Actual system proven through successful operations	Actual application of the technology in its final form and under mission conditions, such as those encountered in operational test and evaluation. Examples include using the system under operational mission conditions.

Table 5. *Technology > Readiness: Technology readiness level index.*

Application	Technology Readiness Level	Definition	Description	Results from
Pre-existing knowledge	1	Basic principles observed and reported	Lowest level of technology readiness. Scientific research begins to be translated into applied research and development (R&D). Examples might include paper studies of a technology's basic properties.	Published research that identifies the principles that underlie this technology. References to who, where, when.
Technology research	2	Technology concept and/or application formulated - conceptual design formulated	Once basic principles are observed, practical applications can be designed (i.e., "invented"). Applications are speculative, and there may be no proof or detailed analysis to support the assumptions. Examples are limited to analytic studies. The step up from TRL 1 to TRL 2 moves the ideas from pure to applied research. Most of the work is analytical or paper studies with the emphasis on understanding the science better. Experimental work is designed to corroborate the basic scientific observations made during TRL 1 work.	Publications or other references that outline the application being considered and that provide analysis to support the concept.

TABLES

Application	Technology Readiness Level	Definition	Description	Results from
Research to prove feasibility	3	Analytical and experimental critical function and/ or characteristic proof of concept - conceptual design tested analytically or experimentally "Concept defined"	Active research and development (R&D). This includes analytical studies and laboratory studies to physically validate the analytical predictions of separate elements of the technology. Examples include components that are not yet integrated or representative. At TRL 3 the work has moved beyond the paper phase to experimental work that verifies that the concept works as expected on simulants. Components of the technology are validated, but there is no attempt to integrate the components into a complete system. Modeling and simulation may be used to complement physical experiments.	Results of laboratory tests performed to measure parameters of interest and comparison to analytical predictions for critical subsystems. References to who, where, and when these tests and comparisons were performed.
Technology development	4	Component and/ or system validation in laboratory environment - critical function/characteristic demonstration "Concept defined"	Basic technological components are integrated to establish that they will work together. This is relatively "low fidelity" compared with the eventual system. TRL 4-6 represent the bridge from scientific research to engineering. TRL 4 is the first step in determining whether the individual components will work together as a system.	System concepts that have been considered and results from testing laboratory-scale. Reference to who did this work and when. Provide an estimate of how test results differ from the expected system performance goals.
Technology development	5	Component and/ or system validation tested in relevant environment "Proof of concept validated"	The basic technological components are integrated with reasonably realistic supporting elements so they can be tested in a simulated environment. The basic technological components are integrated so that the system configuration is similar to (matches) the final application in almost all respects. The major difference between TRL 4 and 5 is the increase in the fidelity of the system and environment to the actual application. The system tested is almost prototypical.	Results from testing are integrated with other supporting elements in a simulated operational environment. How does the "relevant environment" differ from the expected operational environment? How do the test results compare with expectations? What problems, if any, were encountered? Was the system refined to more nearly match the expected system goals?

TABLES

Application	Technology Readiness Level	Definition	Description	Results from
Technology demonstration	6	System/subsystem model or prototype demonstration in relevant environment - prototype/engineering model tested in relevant environment "Proof of concept validated"	Engineering/pilot-scale, similar (prototypical) system validation in relevant environment. Represents a major step up in a technology's demonstrated readiness. Representative model or prototype system, which is well beyond that of TRL 5, is tested in a relevant environment. TRL 6 begins true engineering development of the technology as an operational system. The major difference between TRL 5 and 6 is the step up from laboratory scale to engineering scale and the determination of scaling factors that will enable design of the operating system. The prototype should be capable of performing all the functions that will be required of the operational system. The operating environment for the testing should closely represent the actual operating environment. Examples include testing a prototype in a high-fidelity laboratory environment or in a simulated operational environment.	Engineering-scale models or prototypes are tested in a relevant environment. Results from engineering scale testing of a prototype system that is near the desired configuration in terms of performance, weight, and volume. How did the test environment differ from the operational environment? Who performed the tests? How did the test compare with expectations? What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before moving to the next level?
System integration	7	System prototype demonstration in relevant operational environment - engineering model tested under relevant operational conditions "Demonstrated"	Prototype near (or at) planned operational system. A similar (prototypical) system is demonstrated in a relevant operational environment. Represents a major step up from TRL 6 by requiring demonstration of an actual system prototype in an operational environment.	Results from testing a prototype system in an operational environment. Supporting information includes results from the full-scale testing and analysis of the differences between the test environment, and analysis of what the experimental results mean for the eventual operating system/ environment. Who performed the tests? How did the test compare with expectations? What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before moving to the next level?
	8	Actual system completed through testing in an operational environment "Qualified"	Technology has been proven to work in its final form and under expected conditions. In almost all cases, this TRL represents The technology meets its designed specification.	Results of testing the system in its final configuration under the expected range of environmental conditions in which it will be expected to operate. Assessment of whether it will meet its operational requirements. What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before finalizing the design?
System operation	9	Actual system operated successfully through expected conditions "Proven"	The technology is in its integrated form (i.e., integrated into a service system) and operating under the full range of expected conditions, such as those encountered in operational test and evaluation (OT&E).	OT&E data.

TABLES

Table 6. Technology > Readiness: *Data requirements for each technology readiness level.*

Criteria	Description
Structural servicability	Natural forces, strength properties
Fire safety	Fire resistance, flame spread, smoke development, toxicity, fuel load, combustibility
Habitability	Thermal properties, acoustic properties, water permeability, optical properties, hyegine, comfort, safety
Durability	Resistance to weear, weathering, adhesion of coatings, dimensional stability, mechanical properties, rheological properties
Practicality	Transport, storage on site, handling at installation, field tolerances, connections
Compatibility	Jointing materials, coatings, galvanic interaction or corrosion resistance
Maintainabilty	Compatibility of coatings, indention and puncture (patching), chemical or physical attack
Environmental impact	Resource consumption at production, life cycle impact
Aesthetics	Visual impact, customizing options, color selection, pattern selection, surface selection

Table 7. Technology > Readiness: *Technology readiness levels.*

Applications	Technology Readiness Level	Description	Requirements	Verification	Viability
Broad Range of Applications	1	1.1) Physical Principle	1.2) Needed Capability	1.3) Analytical or Experimental	N.4) Advancement to the Next Level Technical & Programmtics (N = 1-8)
	2	2.1 Basic Concept	2.2) Needed Functionality	2.3) Analytical or Experimental	
	3	3.1) Key Technology Characteristics	3.2) Basic Requirements (Family)	3.3) Simulation or Experimental	
	4	4.1) Full Technology (in the Laboratory)	4.2) Complete Requirements (Narrower Range and Interactions)	4.3) Rigorous Experimental	
Family of applications	5	5.1) Full Technology & Interactions (in a Relevant Environment)	5.2) Complete Requirements (Specific)	5.3) Rigorous Testing at Component and/or Breadboard in Relevant Family of Environment	
	6	6.1) Full Technology in System or Subsystem	6.2) Full Requirements (System or Subsystem)	6.3) Rigorous Testing at System and/or Subsystem in Relevant Environment	
Preliminary Definition for Specific Application	7	7.1) Full Technology in System or Subsystem	7.2) Full Requirements in Space Environment (System or subsystem)	7.3) In Space Demonstration	
Specific Application	8	8.1) Full Technology in System (Manufactured)	8.2) Full System and Qualification Requirements	8.3)Qualification Campaign	
	Application	9.1) Final Manufacturing & Operations Plans	9.2) Performance and Manufacturing Requirements	9.3) System Operations Verification (including life)	9.4) Failure Analysis (if needed) and/or Future

Service Accounting System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: service, service accounting, servicing, service systems, habitat service, service accountability

Abstract

All social-based platforms must account for service. Society is a socio-technical system that must account for service in order for fulfillment to have meaning. Service is an enabling element in society; it enables productive, organized, repeatable, and motivated effort. Service can be accounted for through user and habitat surveys. In the market, service is sold. In the State, service is duty. In a community-type society, services are accounted for, contributed to, and operationalized. Cities are localized service systems. Services are operated by contributors for users. All services require information and objects, and therein, sufficient information and objects to result in a continuation of the service and satisfied users.

Graphical Abstract

Figure 15 on page 71

1 The unified information system

An information system is a fundamental element of a socio-technical society, because it interconnects four fundamental environments: the social and technical spaces as well as the digital (virtual) and material environments, and formalized through signals and language systems allows different actors to interact with coherency and precision. These connections are important in the production of useful projects, designs, possibilities, and simulations that are likely to generate a stable and predictable environment [for human fulfillment]. By viewing society as an information system, it is possible to formalize intentions, perceptions, and physical space in a useful and intelligent manner.

Through a unified information model it is possible to fully account for the material environment, in particular, composition and location. When composition is accounted for, then it is possible can compute various functions of the same materials. With a referential database of materials and functions it is possible to identify probable service configurations - exploring probabilistically the way in which material resources can be transformed into productive goods and services, and then back into their basic material constituents, following a sustainable cycle. Humanity can then design different material configurations of its environment and simulate their engineered experience for optimal resolution of the current habitat.

Everything which has been technically constructed into the habitat may be said to have been engineered and integrated into that which is most often referred to as a "habitat service system".

2 Service system carrying capacity

> **NOTE:** *Population problems have a horrible way of resolving themselves.*

Carrying capacity is a term that relates to the primary service systems in the habitat. Each service system has a capacity determined by its inputs, process, and outputs. The outputs of each service system are calculated to sufficiently fulfill the population, while providing a storage buffer for safety. For example, all cities will be designed with a buffer capacity for housing. Therein, something akin to 3-5% (*an estimate, accurate figures to be determined by decisioning*) of the dwelling will remain unoccupied. This allows for:

1. Expansion of the population,
2. Always available housing alternatives; and
3. Possible emergency housing in case of a disaster.

One might imagine 100% capacity as the most effective strategy for occupation of a locale, but in the context of survival in a larger ecology, a buffering strategy for occupation and usage of a service system is most efficient. Buffering means that there is a lessening or moderating of the impact of something. The buffering part of each service system provides access to resources and other materials in case of an unaccounted for demand or incident. When the precautionary principle is applied to habitat service functionality, then it means to have enough of something so that you have another one if the first one breaks or if more are immediately needed.

Businesses in the market prefer to operate their systems as close to full capacity (i.e., "peak capacity") as possible to maximize their revenues. In community, we design for service and ecological capacity, and we operate within that set capacity threshold with a buffer for risk. In community, there is no incentive to operate at peak capacity. Instead, service operations in community fluctuate directly with demand and participation - they are designed by the user, for the user.

It is also relevant to point out here that populations may actually begin migration within the city network, which may seasonally shift the population sizes of various community-cities.

Any service system may be reconfigured for a new function and capacity.

Allowing expansion sounds like a contradiction to total city design. We can duplicate cities, but to have them undergo expansion may be poor design and not even possible in a sustainable system.

In order to determine a structure's functional capacity, the following must be known (What is the functional capacity of a structure?):

- What is the material composition of a structure?
- What is its engineered configuration?

- What is the functionality that it encodes?
- What are the structure's interdependencies?
- What is the affect of the structure on its environment?

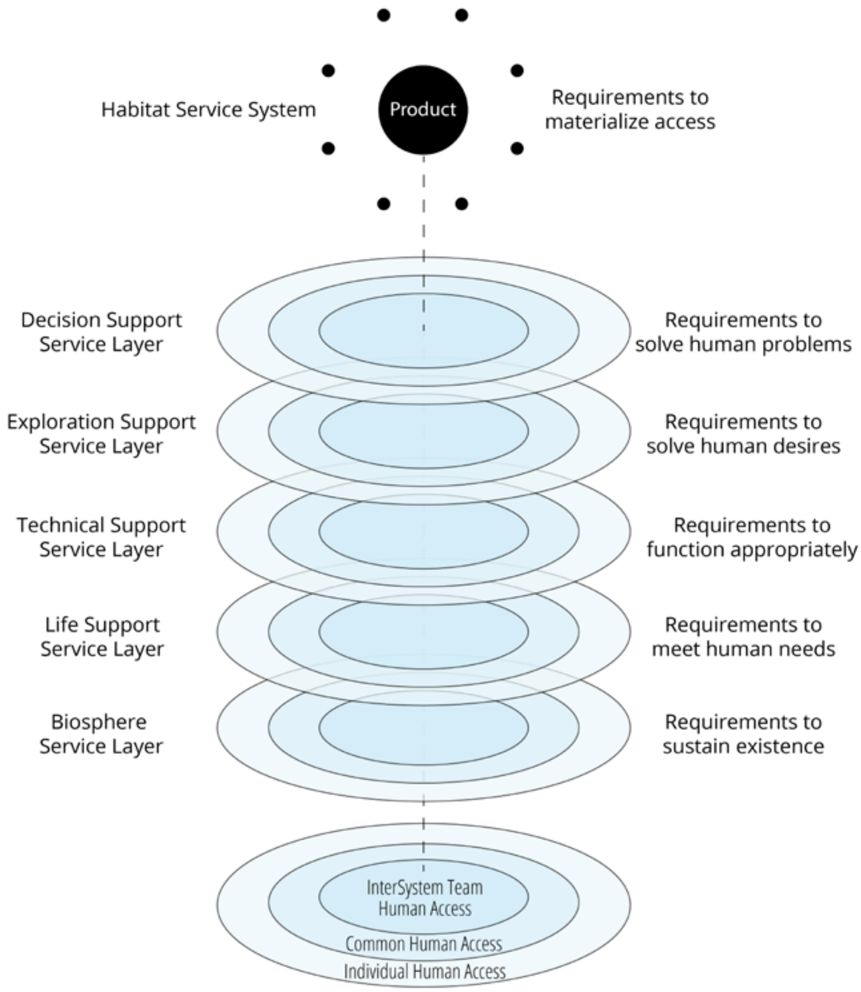
3 Common city services

Common city services found in most habitat service systems throughout community include, but are not limited to:

> *Note that this list is highly limited in detail and scope. This is more of an overview of the services that are available and more fully described in the service specific sections later in this document. Currently, this list acts like more of a comparison between common services in community and "services" the market-State (the second list below).*

1. **Residential services**
2. Accommodation
3. Eating and drinking
4. Clothing and accessories
5. Household and office supplies
6. **Cultivation**

Figure 15. *The service layering of a unified societal system.*



7. Food cultivation
8. Materials cultivation
9. Gardening
10. **Medical services**
11. Emergency medical care
12. Medical safety services
13. Medicines production
14. Life stage care
15. **Energy and power services**
16. Heating, ventilation and cooling (HVAC)
17. **Construction/materialization services**
18. Production (including extractive)
19. Recycling
20. Maintenance and repair
21. **Information services**
22. Computing services
23. Information storage and processing
24. **Communications services**
25. **Transport services**
26. Rapid transport (vehicles, stations, and infrastructure; road and rail)
27. Personal transport
28. Walking
29. Water
30. Distribution
31. Storage
32. **Recreational services**
33. Bodies of water
34. Landscape features
35. Outdoor pursuits
36. Sports and entertainment
37. Sports complex
38. Venues, stage and screen
39. Quite and contemplation pursuits
40. Botanical and zoological cultivation attractions
41. **Science and research services**
42. **Technological development services**
43. **Learning services**
44. Team mentoring
45. Learning facilitation

Services in the market-State that are absent in community include, but are not limited to:

1. **Commercial services**
2. Consultancies
3. Contract services
4. Employment and career agencies
5. Hire services
6. Advertising and marketing
7. Legal and financial services
8. Property services
9. Gambling
10. **Schooling services**
11. **Retail services**
12. **Government services**
13. Policing service

TABLES

Table 8. Service Accounting: *Material habitat cycling human design solution.*

Natural and induced environment factors	Factors that create the physical environment surrounding a team	Human needs of the system	Interface design between human and system
Noise level	Decor	Mutual [threshold] existence (availability = mutual level of availability)	Identifiability
Felt uplift level	Aesthetic level	Availability of sensory stimulation	Information systems and decision support
Vibration level	Anthropometric accommodations	Availability of access	Standardization
Humidity level	Habitat volume	Availability of operation	Re-materialization
Illumination level	Location and orientation aids	Availability of operation	Recycling and waste management
CO2	Transportation and communication paths	Availability of sensory stimulation	Control panels, input devices
Air flow	Fabrication systems and material cycling systems	Availability of information and physical inter-access	Information technology tool availability
Oxygen level (and level of other atmospheric particulate matter)	Windows	Availability of life support	Hardware, software, and InterSystem Team contribution support
Aromatic sensation; Odor	Illumination (as ambient lighting)	Availability of decision support	Availability of InterSystem Team surface area
Toxic: Substances level		Availability of technology support	Availability of Common Access surface area
Atmospheric pressure level		Availability of exploration support	Availability of Personal Access surface area
Radiation exposure level, etc.			Human habitat integration
			Mobility and constraints availability and quality
			Caution and warning functionality
			Human and/or robotics and automation integration

The Habitat Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: habitat service system, city system, cybernetic habitat, cybernetic city, societal service architecture, ecological fulfillment

Abstract

The term 'habitat service system' is the technical name of a 'city'. Cities are human habitats that localize services in a coordinated manner. A habitat can be designed, engineered, operated, evaluated, simulated, and recorded. For every current habitat state, there is a past state that may or may not have been recorded, and there is a future state depicted by the solution to some decision inquiry. Habitat services distribute informational and spatial services and objects to a population. In this sense, a city (localized habitat) is a service bus that connects all interacting functional services in a local environment to one another. A network of habitat service systems (network of cities) may share information and spatial resources in order to optimize service globally. The three core services a habitat service system can offer are life support, technological support, and exploratory support. These three service systems (as well as a biospheric) service form the foundational function of any habitat service system. Each individually localized habitat service system (city) has these three functional service systems, which are engineered into operation through an intersystem [habitat service] team. Contributing team members have intersystem access. The whole population has access via cities to these three core human life functions throughout the community network of cities. Each core functional service (life, technological, and exploratory) has a set of functional sub-service access systems from which users in the community access those services (and service-objects therein).

Graphical Abstract

Figure 16 on page 77

1 Introduction

A.k.a., The city system, the city operation architecture.

The material elements of a society exists within the material, physical environment. The location(s) where humans live and operate within this environment is referred to as a 'habitat'. A habitat (which is Latin for "it inhabits") is an ecological or environmental area that is inhabited by a particular species of animal, plant, or other type of organism. It is the natural environment in which an organism lives, or the physical environment that surrounds (influences and is utilized by) a species population 'habitat'. A habitat sustains a social population through the encoded recognition of a reciprocal interchange between that population and its material environmental reality.

A material environment can be restructured into "intentional" service environments. In other words, out of a common material environment, humans may cooperatively create an intentional habitat to service their common needs. The intentional output of these services systems is: freely and openly accessible services, goods and technical productions (i.e., "products").

Each specific habitat service system (or "service platform") acts as an organizational resource for the structured flow of energy and information (resources) into systems that by their very structure generate a higher potential state of existence within a commonly known environment. The habitat service systems structurally organize common resources toward the fulfillment of individual needs. It could be said that the habitat service system is a platform for the transforming of energy and information into a state that has a higher potential to "support a purpose" and "fulfill a need" [in response].

Herein, operational processes constitute the core functions of these systems and they represent the primary "value stream" (i.e., the end-to-end system process which delivers a service or "product" to an person, subject, or entity). A value stream is composed of a sequence of activities (and tasks) required to design, produce, distribute, and maintain a specific service, with all relevant accompanying information, materials, and knowingly desired conditions (i.e., values).

The habitat service system model represents the functional model of a city. Therein, functions can be defined as the abstracted behavior of a city. Functions are described in terms of the logical flow of information, energy, materials and signals. Functions and sub-functions can correspond to well-defined basic operations on well-defined flows leading to a taxonomy of functions (*as described below*). The functional structure (or, functional architecture) of a city is a form of a conceptual model of the functional domain. A conceptual model of the functional domain is a qualitative representation of the physical behavior of the informational and physical (spatial) structuring of a city as well as the [global] city network within which any city resides. Therein, the physical structure in interaction with a physical environment gives rise to a city's behavior. Behaviors are related to structural-physical descriptions of a city. Behaviors are derived from city functions and their interaction with a material environment. (Stepandic, 2019)

The habitat service system conceptualizes and models the city as a series of homogeneous (Read: alike) and sorted layers, structured around the set of domains representational of human life; that of life support, technology support, and exploratory support. Categorization and taxonomy are important here, as the resulting model seeks functional simplification. These layers are composed of relatively homogeneous, sorted and ordered components, the product of earlier phases of sorting and cataloguing of human life [without the market or State]. Each layer is configured and sorted according to a particular function, that of life, technology, and exploration. Each of its layers is an articulation of a specific logic.

Here, the habitat service system (Read: city) operates through connected classification and taxonomy, not only providing an order but, beyond that, establishing an ontology: categories, attributes and subcategories are created and, in doing so, they create their very object of intervention. Here, reality is thought of as an integrated organizational language and applied stack-a popular way of conceptualizing protocols, data formats and software amongst engineers-ensures that each layer [of the stack] handles the same base information simultaneously, but at different levels of abstraction. Extrapolating 'stack thinking' to the city means that, in a highly hierarchical fashion, different urban systems (such as health, transport, energy or waste) are modelled and understood in the same way (Read: are operationalized together). (Marvin, 2017: 95)

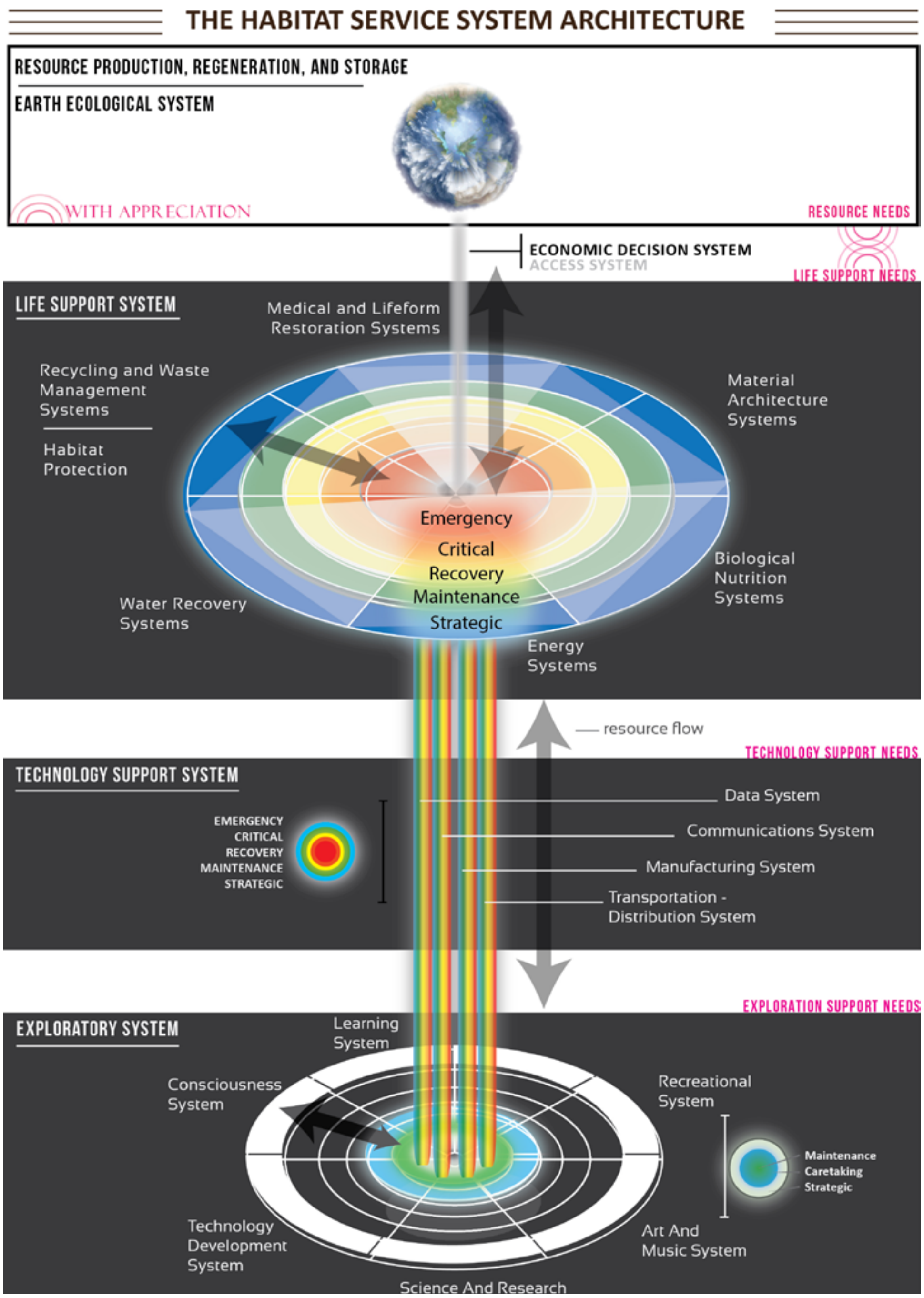
The city is, in essence, subject to a form of modularization and categorization according to a set of predefined [human and ecological] criteria that are then reflected in the realization of a global habitat software and hardware (hybrid) system. In order to integrate city organization, standardization, modularization and classification are fundamental processes. Therein, city planning analysis is the process of breaking down the city into a multiplicity of objects and components.

A service bus is a scheme used in computing, software, and spacecraft development to refer to a transferring interface between mutually interacting components. There are two core service buses to the city, one an information bus (with a particular focus on decisioning) and the other a material [service] bus. These buses represent the core, center or platform around which the wider ecosystem is organized. Within the total ecology of the city there is a form of interlayering of networks, interfaces and data integration that are assembled and operated together in a [decision] control system positioned within the layers of the city. Thus, the city may be viewed, decomposed, as a series of event rules, a set of semantic models, and a set of work-flows that are supported by indicators, directives, and alerts.

The information system of the city uses analytics (data analytics, predictive systems, modelling and simulation) that are based on a set of societal standards for habitat service systems. The analytics generated by habitat data are then related to a set of visualizations, such as dashboards for current operations and future possible operations (i.e., planning). Data integration and gateways for flow control occur along a information service bus and within the information system itself, which brings into existence a real-time, real-world visualization (Read: model) of the operation (or, potential operation) of the system. This visualization can be viewed from several core perspectives, including that of the support services themselves, the software therein, and the hardware therein. Such a holistic view of the habitat as an integrated information and spatial (Read: material) system, where everything is a data point, allows for flexibility, efficiency, and optimization of the planning and operation of the environment for all inhabitants.

Within the city (Read: habitat service system) network, there is the ability to access data globally, as well as the need for modularity, interoperability, and transferability across [service] systems and cities. An yet, each city within the network is also a customized package of sub-services (or, sub-customizations of service) depending upon the unique circumstances of individual cities. Local issues enter the global city information network in the form of data. Therein, by combining data sets, cities may

Figure 16. *The Habitat Service System Decomposed Layered Reference Model.*



be reconfigured in a multiplicity of ways. Therein, cities maintain a central processing system (or, central processing unit, CPU) as part of their information support service, which processes not only local city information, but distributed information pertaining to the global city network, which from an information viewpoint is known as, the societal information system. The societal information system works on comprehensive design solutions that may be applied to any city in the network. This process of disaggregation is made possible by reconfiguring the components of the city into data blocks that can later on be worked with, recombined or reprocessed. The city is viewed, like the society itself, as an information system (an assemblage of data), which may be disassembled into its constituent parts as defined by the categories of any human-based habitat service system, and then unproblematically re-assembled into new more desirable configurations and flows. Therein, [habitat service] operational processes can be analyzed as data packages and reconfigured in a variety of custom ways. (Marvin, 2017: 98-99) This technique is sometimes given the term 'digitalization'; and, the logical computation of a digital (information) system for a city/habitat is often called, 'habitat computational logic' (a.k.a., city computational logic). Whereupon, the total logic of said environment for a operating system for the global habitat (a.k.a., habitat operating system or city operating system. In a global, technologically developed community-type society, computational logics have become ubiquitous, pervading every aspect of life.

> **INSIGHT:** *A comprehensive habitat systems approach recognizes that the fabric of the natural world, from human biology to the Earthly biosphere, to the electromagnetically gravitational arrangement of the universe itself, is one huge synergistically connected system, fully interlinked. Human cells connect to form organs, organs connect to form bodies, and since bodies cannot live without the Earthly resources of food, air, water and shelter, organisms are intrinsically connected to the Earth in each moment of breath.*

1.1 Societal access platforms

All societal-based platforms must account for a material system. When producing anything, access to objects must be accounted for. Access is necessary and two dimensional concept. Firstly, there is access to a team or working group through a contribution-based structure, and then, there is access to goods and service (without force of trade). Access can be accounted for many types of surveys including demand surveys, resource surveys, contribution surveys, etc. In the market, access is considered through the cost of a sale. In the State, access is acquired through authority. Humans require access to objects and information, which are composed into services. In a market, access is controlled by price, and the concept itself is mixed with "rights" (given by authority) and "property" (purchased in the market). In a community-type society, access refers to demands and other issues for service that are accessible to users. Ultimately, the goal is to have access to that which optimally meets user requirements (human needs) given that which is available at the time of access. In a community-type society, access centers and integrated transportation systems distribute products. Services are integrated, often modularly, into the infrastructure of the environment in order to optimize efficiency and produce a higher quality experience of access [to services] by a user. With sufficient technical knowledge and ability it is possible to apply automation technologies to increase the efficiency by which access occurs. Automation technologies can free individuals for access to opportunities they might otherwise not have had. Automation technologies can also make access to services, such as medical and informational more safe, reliable, and faster.

1.1 The habitat system states

> **NOTE:** *In nature, a 'structure' is a responding service. And herein also, the habitat service system is structurally designed as responding services (i.e., a service that responds appropriately to human need).*

In systems thinking the state of a system is a complete description of the system in terms of its condition, its parameters, its dynamics, values and variables, at a particular moment in time. This domain represents the formalized, existent structure of the community (the one actually operating or previously operating).

The Real World Community information system maintains a record of every known state of every system in the habitat. This includes both a model of the natural world, and a 'state model' of each service system.

There are three possible types of state for which the information system must account:

1. The **current state** of each habitat system (*quantitative* and *qualitative*).
2. The **past states** of each system of the habitat are identified as elements of the habitat's history (*quantitative* and *qualitative*).

3. The **future planned, predicted, and simulated states** that identify potential states as well as the next selected incremental state (*probabilistic*).

The 'past' represents a record of former re-structured iterations of the environmental habitat. A 'past state' represents a model of the prior state-dynamics of information, energy and services in our total environment.

The 'current state' space represents the current re-structured iteration of our environmental habitat -- the current state dynamic of information, energy and services (Read: the responding flow of resources) in the our total environment.

Individuals in community naturally seek the iterative improvement of their service system's trajectory toward greater states of human fulfillment. In other words, in community, our intention is to cooperatively create progressively more informed and fulfilling states of our habitat.

NOTE: *It is useful to know where we have been so that we can intelligently design where we are going. Further, it is useful to simulate where we are going so that our likelihood of a safe arrival is more certain.*

2 The service-oriented architecture of a global habitat service system

INSIGHT: *In any architecture, energy can be spatially and temporally positioned within that architecture in a variety of ways. For example, oil, coal, natural gas, and nuclear are highly centralized providers of energy [as electricity]. Solar, wind, and to a lesser extent hydro, geo-thermal, and biomass, can be localised and provide the energy requirements of a community that seeks electrical generation at a distributed level. In either case, the energy derived therefrom could be laterally decentralized [in time and space] into a series of backup batteries.*

The Habitat Service System is an integrated system for servicing the fulfillment of the material needs, wants, and preferences of individuals in the community. This type of an integrated service system is also sometimes known as an "functionally integrated city system"; yet, it might be otherwise referred to as an "functionally integrated habitat system". It is designed as a total "functional service platform" for the community in harmony with nature, existing within and through the habitat -- it is a part of the ecological habitat that we have formally and technologically redesigned to service our needs in a manner that is technically functional and commonly fulfilling.

When a group of people are living within the same community and sharing resources, the systems that support their lives together must be identified, operated, and optimized for the community's very survival.

A basic consideration in the design of habitat service systems is that of dividing work (as effort and services) into reasonable and prioritized tasks and activities (as time and spatial differentiation), while giving simultaneous attention to coordinating these activities and unifying their organization into a meaningful whole (as integration) so that the system can adapt and re-orient where said response is desired.

The Habitat Service System is principally divided into four service sub-systems. These systems are connected to one another within the larger and more encompassing Real World Community information model. In their layered portrayal, they are seen with the decision system running through each of their layers. Each service architecture functions to fulfill a particular category of need (in a temporal and spatial manner). Each system in the habitat involves the nesting of subsystems that must operate together for the overall system to work effectively.

The four global habitat service systems are:

1. **The Resource Production, Regeneration and Storage System (planetary biosphere)** - provides for the community's resource needs - the natural

phenomenological environment, the planetary lifeground. Strategic preservation of the lifeground is a requirement for the continuation of all other service support systems. This is the planetary system itself.

2. **The Life Support System (LSS)** - provides for the community's life support needs. This system might be equivocated with the idea of "needs". Provides for material life support functions; the life support platform. This is the life-sustaining platform, including necessary infrastructure, for a population. This could be considered a social infrastructural system with a dedicated life function.
3. **The Technology Support System (TSS)** - provides for the community's technology support needs. This system might be equivocated with the idea of "wants". Proves for technology support functions; the technology support platform. The is the technical infrastructural system for a population.
4. **The Exploratory Support System (ESS)** - provides for the community's exploration (and therein, discovery, self-/social-development, and recreational) needs. This system might be equivocated with the idea of "preferences". Provides for self and social exploration functions; the exploration support platform. This is the self- and social-development platform, including necessary infrastructure, for a population. This could be considered a social infrastructural system with a dedicated exploration function.

Each of these systems represents a functional service, a platform, that has been separated out to meet the [frequency] needs of humanity using resources from the common environment. Essentially, these service systems differentiate the different functions that control the 'phenotypic expression' (to use a term from genetics) of a community. Fundamentally, a functional approach allows for the identification of root concerns and the implementation of systemic solutions.

The primary four functional service systems are each sub-composed of a further set of functional subsystems. These subsystems exist to meet the ongoing and delineated functional requirements of each of the four primary categories of need. These sub-service systems fulfill needs by generating [responsive] access to technical production services. Wherein, for instance, the Life Support System is sub-composed of six systems, each of which transfers energy and resources into a particular category of good or service designed to meet the ongoing functional life support needs of individuals in the community. Essentially, these service sub-systems sustain the functioning of the community and are permanent structural elements of the Habitat system. They exist as long as our need for them exists.

The integrated habitat service sub-systems are:

1. The Life Support System is sub-composed of:
   A. *Architecture System*
   B. *Water System*
   C. *Power System*
   D. *Medical System*
   E. *Cultivation System*
   F. *Recycling And Waste Management System*
   G. *Defense System (non-primary)*
2. The Technology Support System is sub-composed of:
   A. *Information System*
   B. *Production and Recycling System*
   C. *Transportation and Distribution System*
   D. *Communication System*
3. The Exploratory Support System is sub-composed of:
   A. *Science and Research System*
   B. *Technology Development System*
   C. *Learning System*
   D. *Recreational System*
   E. *Consciousness System*
   F. *Art and Music System*

All service systems act independently as well as interdependently - they follow dynamic, distributed systems principles - they are centralized and decentralized. It is inaccurate to label them as centralized *or* decentralized, as one or the other. Most issues involve a spectrum of subsystem requirements, and therefore, necessitate the involvement of distributed multi-system effort (i.e., multiple systems acting together to meet a need or accomplish a purpose).

Each sub-system may be seen not only as an area of service, but also an area of inquiry. As such, the word "science" is sometimes appended to the end of the name of each system. For example, "water recovery sciences" or "biological health sciences", and so forth.

In concern to measurement of the community system as a whole:

1. These functional processing systems are a measure of the technical efficiency of the community.
2. The alignment of these systems with

the community's current understandings and technological development is a measurement (indirectly) of the technological age of the community
3. The functioning of these systems are [in part] a measure of the technical resiliency and health of the community.

Issues of greater urgency and those of a strategic nature are more likely to involve multiple system interdependencies, and are sought resolution through an interdisciplinary systems approach (i.e., a systematic solution orientation). In particular, "urgent" issues have the potential of impacting the stability of service systems, and therefore, they require rapid response and a high degree of systems-level coordination. Similarly, "strategic issues" involve the planning of future states of a the total habitat system, and therefore, coordination among systems is relevant. "Operations and maintenance" issues assume a more direct and targeted approach by individual sub-systems, and they have fewer interdependencies; although, it does occur that some maintenance issues involve multiple systems.

The Life Support and *Technology Support Systems* represent the Habitat's **core service systems**. The Exploratory Service System is one of the four primary Habitat systems, but it is not a "core system"; it is a secondary system because it relies on outputs of the two core systems to maintain its existence. The Exploratory system exists because:

1. The critical life support needs of individuals in the community are sufficiently fulfilled (i.e., the Life Support System is functionally operational), and
2. The Technology Support systems is functioning at a sufficient threshold to then begin meeting the needs of the Exploratory System (i.e., The Technology Support System is functionally operational).

Thus, exploratory system issues are prioritized after the critical requirements of the Life Support and Technology Support Systems, for if they fail then every system "downstream" will fail also.

Here, functional community design relates an individual organisms resilience to the resiliency of the community as the ability resist illness, the ability to resist injury, the ability repair, the ability to reproduce, to have movement, to generate energy, and to direct energy into a functional state rather than just lose energy to the universe.

The sub-systems of the primary four service systems maintain an operating structure that involves the operational processes of *integration and planning*, *operations* and *maintenance, and incident response*. 'Operational processes' define the primary tasks (or activities) that must be performed to ensure the stability and continuity of the whole Habitat Service System.

The three operational processes are:

1. **Strategic integration and strategic preservation planning** - The process of integrating goals, values and new understandings into the design of future services and technical productions. This operational process involves decision planning. A society with a purpose must have a set of blueprints, as well as a planned set of blueprints. This operational process is also sometimes known as "Strategic Planning and Preservation". The community uses the operational process of Strategic Preservation Planning to ensure that needed goods and services are continuously accessible.
2. **Operations and Maintenance** - The process of preserving and improving the ongoing functioning of the system so that it continues to provide goods and services as planned and as happens to occur. The Maintenance and Operations process ensures the continuation of systems which provide for individual access to products and services. This process approaches the concepts of integrity, availability, and transparency through increasingly efficient and automated action. As a structure, this operational process transforms resources in an effort to maintain, and improve the quality of access and use of resources, goods and services.
3. **Incident Response** - The process of responding to malfunctions and other [potentially harmful] incidents within the habitat system to maintain the urgent homoeostasis (i.e., critical self-regulation) of the system. Incident response is essentially the critical resolution of a point of identified failure in the system. Highly reactive issues are urgent and they have the potential to impact the integrity, availability and transparency of systems. This operational process includes processes involved in the recovery from malfunctions and other incidents.
   - **Fail-safe** and **fail-secure** (as a task for safe systems engineering and redundancy planning) - in the event of failure, the system responds in a way that will cause no harm, or at least a minimum of harm, to other systems or danger to individuals. Fail-safe means that a device will not endanger lives or other systems when it fails. A system's being "fail-safe" means not that failure is impossible/improbable, but rather that the system's design prevents or mitigates unsafe consequences of the system's failure.

Table 9. Habitat*: Protocols for habitat coordination and control operations.*

Access Service Control Types	Control protocols
Resource service control	Resource Accounting
Production service control	Strategic preservation
	Strategic safety
	Strategic efficiency
Demand and Distribution service control	Strategic proximity

Each habitat service support system is composed of the same three operational processes. The operational processes generate actions that provide for the community's purpose, orientation, requirements, and needs. A system process can be decomposed into several sub-processes, which have their own attributes, but also contribute to achieving the goal or purpose of the super-process. The analysis of system processes typically includes the mapping of processes and sub-processes down to an activity and task level, including a description of the "constructor entities" as well. Each operational process can be subdivided into its base activity and task level, which are interrelated within a comprehensive, real world information system.

These systems-level tasks are planned for through a decision system, and carried out by [inter-]systems teams who have demonstrated experience, or are being mentored by those who have demonstrated experience.

> **NOTE:** *The operation of a service function necessarily takes up material space and time.*

2.1 The Resource Production, Regeneration And Storage System

> *A.k.a., The natural environmental domain, the world, the planet, the life-ground, the Earth's ecological system, the ecology, the biosphere.*

Necessarily, a society must construct in, and account for, its environment, continuously. The Resource Production, Regeneration, and Storage System is the natural [planetary and solar] environment; it is the world that creates (or "has created") all of the resources humanity has access to; it provides for humanity's resource needs, including the production, regeneration, recycling, and storage of resources. The natural environment is the material basis for human survival and socio-economic (socio-technical) development; it is the environment from which humanity acquires resources, discovers knowledge, and into which the material systems of a society (i.e., the habitat service systems) are produced and integrated. Fundamentally, the natural world provides for humanity's resource needs and life experience.

The planet is ultimately where all resources that humanity has temporarily accessed return to, after they are used, or when access is no longer required. The products that humanity produces from the planet's resources will eventually decay and be recycled. Physical life requires resources from the environment for its development and continuance. Therein, there is no life without death.

The Earth and the services that it provides represent a common [life]ground for all of humankind, and all present symbiotic life. Hence, to simply treat the natural environment as a physics lab is folly, potentially beyond repair, and is the ultimate form of irresponsibility. Planetary resources are finite, and it is important to be responsible (and efficient) in their use.

Technological systems can facilitate and optimize this system. For example, a building can store harvested resources, and production technologies can increase the nutrition content and yield of a harvest.

2.2 The Life Support System (LSS)

> *A.k.a., Life support service system, life sustaining system, life system, environmental control system and life support (ECLSS).*

This Life Support System exists to meet the functional *life support needs* of the community. The life support system is further divided into subsystems representing the essential service categories (or functional processing categories) for the direct support of human life within a habitat. Effectively, life support refers to the vital service functions (and their outputs) for sustaining human life. These systems provide services and products for which everyone in the community has a life-need. In other words, everyone in the community has a direct bio-psycho-social primal need for the outputs, and other technically service productions available *through* each of these subsystems. Each one of these core functional processes (or subsystems) is required for life stability and is a possible point of resiliency failure for a habitat (i.e., is somewhere that a community can fall out of the state of resiliency).

Individuals will always have a need for food, water and shelter. They will always have a need for the production of energy. They will always have a need for medical care and the recycling of waste in their ecology. These life support needs are critically common. Every habitat service system needs at least these systems on a continuous basis - these are components of a core habitat service system.

The Life Support System maintains the idea of 'social assurance' as the basis of community resiliency and true economic "security" -- that the systems that compose the habitat may be accessed in such a way that the lifeground is preserved and humanity's life needs are fulfilled.

The subsystems of the Life Support System are:

1. **Architectural service system** (a.k.a., building and clothing service)
    - All the activities and objects associated with architectural buildings and clothing, including but

not limited to: design, storage, transportation, and usage.

2. **Water service system** (a.k.a., hydrology service, water cycling service)
   - All the activities and objects associated with water, including but not limited to: storage, transportation, recovery, processing, and usage. The water service system includes other liquids and atmospherics.
3. **Power service system** (a.k.a., energy service, energy-power system)
   - All the activities and objects associated with power, including but not limited to: production, transportation, energy storage, and usage.
4. **Cultivation service system** (a.k.a., nutrition and textile service)
   - All the activities and objects associated with cultivation of food and organic textiles, including but not limited to: cultivation, storage, transportation, and usage.
5. **Medical service system** (a.k.a., life-form restoration)
   - All the activities and objects associated with medical, including but not limited to production, storage, usage, and procedures.
6. **Defense service system**
   - All the activities and objects associated with defense of the habitat.

Note that because the whole habitat system is not yet complete, there are currently several different views on the specific sub-composition of the Life Support Service System. These differences are:

1. Defense as a core life support service, or defense not included at all.
2. Recycling and waste management as a core life support service, or the placement of this system with the Technology Support Service System as part of the Production System (a.k.a., Materialization System).

2.2.1 Survival

In the wild, in a true survival-based situation, there is an ordering to human efforts toward need fulfillment. The prioritization in a survival situation is:

1. Shelter
2. Water
3. Fire
4. Food

These are the original four survival needs for which humanity can produce technologies that function toward improving those conditions necessary for survival.

> **NOTE:** *Many in early 21st century society have become ignorant to what it takes to survive and thrive together on this planet. Our actions, at the incremental level, can generate a greater likelihood of suffering for all others. Today we are becoming far more aware and realize that we are all connected in our lifestyles and materializations on this planet. We are all connected; the boundaries that we may perceive do not exist. We need symbiotic relationships, particularly between our individual selves.*

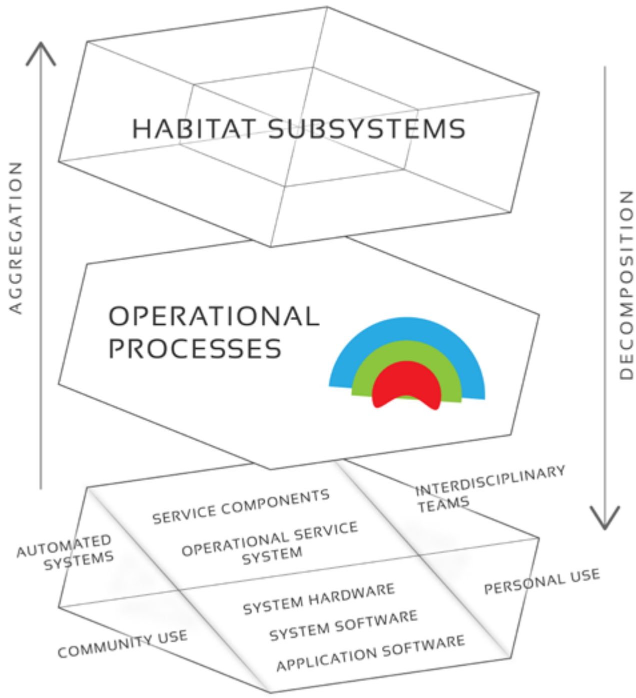
Of note, the primary functions of shelter are:

1. One of the primary functions of a shelter is to enable you to maintain a homeodynamic body temperature.
2. The buildings we construct, which provide an environmentally controllable space
3. Clothing, including shoes and hats, is a form of sheltering.

2.2.2 Defense

Defense is the expression of force to counteract incoming force. As with any defensive system it can be

Figure 17. *High-level aggregation/decomposition layering of the Habitat Service Support System.*



your best friend or your worst enemy. Our biological immune system is a great example of this: if there is nothing to defend against and the defense system wants to be active, then it might start attacking things around it (e.g., the modern military-industrial-prison complex). In medicine, such behavior is casually called "innocent bystander activity" wherein the immune system begins creating inflammatory diseases like autoimmunity, allergies, and neuropathy -- it is trying to do battle against a feigned enemy that doesn't really exist. There are three interrelated ways to prevent the triggering of self-/social-harm:

1. By redesigning the structure of the system with an improved understanding of how to limit the regeneration of conflict between engineered structures and other structures in the ecological habitat [by learning from mistakes and correcting].
2. By redesigning the structure so the triggers of conflict are not present.
3. By removing the "offender".
4. By "cooling" the system (i.e., behaving in a way that avoids inflammation while stimulating healthy behaviors). In other words, the defense system of our habitat service architecture needs to be of a particular [emergent] structural design so that we aren't unwittingly harmed by it.
5. By facilitating the evolution of consciousness through the techniques of self-development.
6. By individually releasing trauma.

2.3 The Technology Support System (TSS)

A.k.a., Technology support service system.

The Technology Support System functions to meet the *technology support needs* of the community. Technology is the application of scientific knowledge for practical, socially identifiable purposes. The Technology Service Support System acts as a conduit for information, energy and resources as they move through habitat. The technical optimization of their flow generates a greater potential for the extension of ourselves into our environment (i.e., they extend our functions). Both individuals as well as other systems in the Habitat System have a need for technological support services.

> **INSIGHT:** *What good is technology if a society does not have the wisdom to use it to better itself and enhance the lives of everyone? Anything less than this will simply lead to a dysfunctional, technologically dangerous society.*

The subsystems of the Technology Support System are:

1. **The information system** (a.k.a., information storage and processing)
   - All the activities and technologies required to process information including computing technology and software systems. This system involves information processes.
2. **The communications system**
   - All the activities and technologies required to communicate including communications technology and software systems. This system involves communications processes.
3. **The transportation system** (a.k.a., transportation and distribution, transport and material distribution)
   - All the activities and technologies required to transport and distribute materials. This system involves logistical processes.
4. **The materialization/production system** (a.k.a., production and recycling service system, materialization service system, materials cycling, technical production and recycling, materials cycling and waste management)
   - All the activities required to acquire materials, produce technologies, and cycle materials, including. This system involves materialization processes.

Solid waste results from the following:

1. Metabolic output of humans and other organisms.
2. Food preparation
3. Material primary system operation (e.g., production, packaging, etc.)
4. Material subsystem operation (e.g., residual substances from water processing)

Firstly , the waste is collected and, if necessary, separated. According to its composition and the applied concepts for disposal, it is shredded, compressed, chemically and biologically stabilized, and then stored.

> **NOTE:** *A "facility" is designed to afford the function of a convenience or service. The term, 'facility', comes from the French language word "facile", which means easiness or with ease. This model defines 'facilities' as - those systems that make life easier, more liveable, and support the community's highest development by meeting the quality-of-life [social and recreational] needs of a community.*

2.4 The Exploratory Support System (ESS)

A.k.a., Exploratory support service system, exploration support system.

The Exploratory Support System functions to meet the exploration requirements of the population. Herein, exploration includes, but is not limited to discovery, expression, and self-development activities. The Exploratory Support System is aimed at providing the services and products to facilitate exploration of the world and of one's own higher potentials.

> **CLARIFICATION:** *Exploration is the act (or actions) of searching, discovering, developing, and/or expressing.*

Humans have desires beyond basic needs. If this were not true then there would be no inventors, designers, no exploration and creativity. The Life Support and Technology Support Systems together allow for the stable existence of the Exploratory System. Although the Exploratory System is a separate system, it relies in great part on services from the Technology and Life Support Systems to operate.

The Exploratory Support System is composed at a high-level of the following sub-systems:

- **Science and Research System** (a.k.a., scientific research and engineering development)
- **Technology Development System**
- **Learning System**
- **Recreational System**
- **Art and Music System**
- **Consciousness System**

> **NOTE:** *The term "sciences" could be added to the names of these subsystems: the learning sciences system; the recreational sciences system; the art and music sciences system, the learning sciences system; the research sciences system, the technology development sciences system, and the consciousness sciences system.*

Scholarly references

- Alling, A., Thillo, M.V., Dempster, W., et al. (2005). *Lessons learned from biosphere 2 and laboratory biosphere closed systems experiments for the mars on earth project.* Biological Sciences in Space, 18(4), pp250-260. [jstage.jst.go.jp]
- Marvin, S., Luque-Ayala, A. (2017). *Urban Operating Systems: Diagramming the City*. International Journal Of Urban And Regional Research. DOI:10.1111/1468-2427.12479 [onlinelibrary.wiley.com]
- Messerschmid, E., & Bertrand, R. (1999). *Environmental Control and Life Support System*. Space Stations, 109–145. doi:10.1007/978-3-662-03974-8_4
- Ruck, T., Putz, D. (2019). *Dynamic Simulation of Performance and Mass, Power, and Volume prediction of an Algal Life Support System*. Technical University of Munich, 49th International Conference on Environmental Systems. [ttu-ir.tdl.org]

TABLES

Table 10. Habitat Service System > SubSystems*: Habitat service system tiers.*

First Tier System	Second Tier Systems (Subsystems)	Third Tier Systems (Subsystems)	Fourth Tier Systems (Operational Processes)	Activities & Tasks
The Habitat System	**Resource Production, Regeneration And Resource Storage; Life Support System; Technology Support System; Facility/ Exploration System**	**Shelter/Architecture; Power/Energy; Nutrition; Water/Atmospherics; Medical; Recycling & Waste Management; Data Processing; Communications; Manufacturing; Transportation & Distribution; Recreational; Art & Music; Science And Research; Technology Development; Consciousness; Learning**	**Strategic Planning And Preservation; Operations & Maintenance; Incident Response**	**Not Identified In This Table**

Table 11. Habitat Service System > Sectors: *The Habitat Service Systems and their secondary sub-systems. This table layout of the service systems (i.e., their aggregation) allows for, or otherwise facilitates, economic calculation. Life, technology, and exploratory services all have a final user demand. Life and Technology services have an intermediate demand, and two exploratory services of Scientific Discovery and Technology Development, also have an intermediate demand. To have an intermediate demand means to require something necessary for production of the final demand by the user.*

Top-level Habitat Aggregated Service Systems	Secondary-Level Habitat Aggregated Service Systems	Service Platform Tasks	Service Platform Resource Compositions and Allocations
NEEDS	DEMANDS	OPERATIONS	RESOURCES
Life Support Service System	Architectural service	...	...
Life Support Service System	Water service	...	...
Life Support Service System	Cultivation Service	...	...
Life Support Service System	Power Service	...	...
Life Support Service System	Medical Service	...	...
Technology Support Service System	Information Service *(Storage and Processing)*	...	...
Technology Support Service System	Communications Service *(Devices and Protocols)*	...	...
Technology Support Service System	Transportation Service *(Machines and Protocols)*	...	...
Technology Support Service System	Materialization Service *(Machines and Protocols)*	...	...
Exploratory Support Service System	Scientific Discovery Service	...	...
Exploratory Support Service System	Technology Development Service	...	...
Exploratory Support Service System	Learning Service	...	...
Exploratory Support Service System	Recreation Service	...	...
Exploratory Support Service System	Art & Music Service	...	...
Exploratory Support Service System	Consciousness Service	...	...

Life Support: Architecture Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

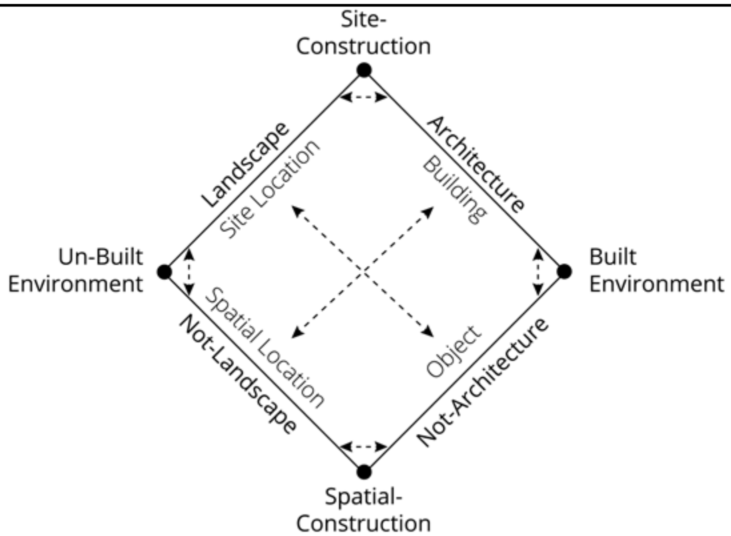
Keywords: architecture, structure, infrastructure, suprastructure, buildings

Abstract

There are several different common names for that which has been engineered to separate humans and their objects from external environmental influences. The name given to the results of this process is varied, and often, discipline dependent. Most simply, the name given to any such deliverable could be "structure" - something created on land or some other platform used to contain ("house") humans and selected objects. The fundamental purpose of any structure is to control the separation of two objects or two concepts. In the process of separation, a pre-existing flow of information and/or matter may be changed. Hence, structural change often conveys functional change, because there is always motion. Other names for a structure include, infrastructure, architecture, buildings, and clothing *(these are all structures)*. Structures can be on a physical individuals person, such as clothing. Structures can be positioned on land and on other surface platforms. Most structures are buildings. Inside building structures are a sub-structure called a room. Buildings (structures) can consist of one or more rooms. Some structures are not buildings. Some structures are not buildings per say, but are containers of objects with useful function (for example, a light post or radar array). Technically these structures are still buildings; they are just buildings designed to "house" technology and not humans. Similarly, there are structures used (interacted with) directly by humans, such as merry-go-rounds and boats. There is also the possibility of structures with purely aesthetic (Read: look and/or feel) function and no ability to contain other objects.

Graphical Abstract

Figure 18. *The architectural representation of a structure on a landscape or other platform for use by humans or having some other function.*



1 Architectural engineering

A.k.a., Structural engineering, infrastructural engineering.

Architecture is the sculpting (i.e., re-configuration) of space. space has power relative to its configuration (i.e., it means "power" and has social influence). Architecture can reconfigure "power" relationships in society. who is excluded and included from the space is essentially the difference between utopia and dystopia. Architecture that uplifts everyone. Architecture is the conceptual design and fundamental operation of a structure. Historically, it was the craft of building shelter.

Architecture refers to the design of structure - the designed structure of something. A 'design' is a visualization of a set of relationships which form a constructable and functionable object or service. In other words, a 'design' is a description of a construction, which may or may not have been constructed. Architecture is a communicable representation of that which has been constructed, is constructed, or could be constructed. In its physical application, architecture becomes the integration of structures, materials, and construction technology.

Architecture is both the process and the product of planning, designing, and constructing buildings and other physical structures. It refers to the materially built environment and necessarily involves mathematics. Every material construction is the result of architecture. And, everything that is a material construction involves architecture (not just buildings). Physical architecture is not limited to buildings. If something involves a relationship with material form, then it involves architecture. Not not all processes, however, are architectural processes. To a large degree, this is why architecture is sometimes known as "the mother of all arts".

Architecture is a life-support functional service, which has historically been given the survival-oriented term "shelter". Your clothes, shoes, and the buildings around you are all part of this "primitive skills" category known as "shelter". In community, that which has historically been known as "shelter" is now called "architecture". The design of structures (i.e., when doing "architecture") involves working with elements that are put together to accomplish a life-orienting function - the use of structure to control for environmental variables in a given space.

Architecture has to do with planning, designing and constructing form, space and ambience to reflect functional, technical, social, environmental and aesthetic considerations. Its realization involves the reconfiguration and coordination of materials and technology. Just like social organization can produce a social "navigational" model, decision organization can produce a decision space resolution algorithm, the product of architectural work is typically drawings, plans and technical specifications, defining the structure and/or behavior of a system that is to be or has been constructed.

> **FOR EXAMPLE,** *architecture concerns the behavior of humans in how they use a space, as well as how the structure of the space relates to its own surrounding.*

When expressed materially, architecture becomes the spatial structures we use and pass time within. Architecture is the physical structures, both permanent and impermanent, which are created into the city for any period of time.

> **INSIGHT:** *In the natural world, adaptive structures are the result of conscious self-organization.*

Any society is capable of being identified by its architectural constructions. However, such a label would be an imprecise representation of the totality of a given society because it doesn't account for the societal living system as a whole, which necessarily includes a social organization, a decisioning organization, and an expressed lifestyle.

In the controlled habitat of the Community (i.e., the Habitat Service System, the "city"), everything is analyzed and broken down for functional use by its users.

> **NOTE:** *Among community there is always a purpose to [architectural] presence, where architecture becomes an emergent completion of nature.*

When it comes to architecture there are rules. There are rules because you don't want a roof caving in on you or your clothing failing when you are at risk to exposure.

> **QUESTIONS:** *How do we scientifically respond to architecture, and how we might best design our architectural environments to fulfill our needs?*

1.1 Clarification of the term 'architecture'

In a material environment the term 'architecture' has several related meanings, which together relate to the science of designing and organizing physical structures:

1. Architecture as an organized structure for implementing a technology.
2. Architecture as an organized structure for reifying a concept.
3. Architecture as the science of designing buildings (and other physical structures).

1.2 Engineering

Architecture is an operational service process accounted for in design and development engineering. Engineering is a knowledge application process for designing and developing services, including architectural services, nutrition services,

Engineering is the application of mathematics, empirical evidence and scientific, economic, social, and practical knowledge in order to invent, design, build, maintain, research, and improve structures, machines, tools, systems, components, materials, and processes. All architecture is engineered. Engineering involves the application scientific principles to a design to ensure it functions as intended. Engineering is part of the solution inquiry process of the Community's common decision space. Matter, which is the expression of nature we pickup signals from and interact with via this physical vehicle, is designed (i.e., architected) and engineered into different configurations for our required fulfillment.

One could state that "architectural engineering" is the application of engineering principles and technology to architectural design and construction.

To understand the interrelationship between architecture and engineering in community design it is important not to have your thinking limited by the language of a paradigm where all economic work is segregated into labor disciplines like "mechanical engineering", "architecture", and "industrial engineering". Instead, it is important to think in terms of systems and the ecologies of systems. All architectural design requires engineering, and we engineer everything that comes into material service as part of the habitat service system, which exists in an ecological dynamic with a larger environment.

> **NOTE:** *In engineering, adventurousness is akin to incompetence.*

1.3 Architecture reflects consciousness and directs its experience

> **APHORISM:** *We first build the roads and then the roads build us. We first make the house and then the house makes us.*

Architecture can facilitate or get in the way, of us experiencing the fullest from our environment. We don't want to put up façades that block our direct experience. Our creations can just as easily trap things in, as keep things out.

Architecture is, in a way, a reflection of those people who have manifested it and maintain its construction. Materializations (e.g., buildings, landscapes, and technologies) are, in a way, like reflections of the people who manifest them. Their design and appearance give an indication of the occupiers' characteristics and understandings. They are a reflection of their integration and realisation as well as their individuality. The materializations of a population are a chosen representation of its character.

> **INSIGHT:** *Momentum represents forward movement. Monuments represent movement in the past. A current theory is like a monument made to an old fact. Most of early 21st century society has lost all momentum and has become a monument to old thoughts.*

Architecture can shape us (our thinking and behaviors) in ways that we don't realize, and yet, are highly predictable.

2 Atmospheric sightlines

A.k.a., Sight line, visual axis.

A sightline is an unobstructed line of sight through the atmosphere between an intended observer (or spectator) and a subject of interest, such as a building, area, etc.

INSIGHT: *Buildings turn into "sculptures" when seen from a distance.*

2.1 Architectural aesthetic pollution

Early 21st century sightlines are highly polluted with human [mental and commercial] "defecant". The question must be asked, Why are people "defecating" in their visual environment? For their very well-being, people do not defecate in the same water they drink from. Why then are they expelling into their visual environment, whatever is to their mind, without consideration given to the larger environment and its many users?

3 Starchitects

Modern cities are monuments to "great names" and amusing attractions, most of which are not representative of humankind's true potential for fulfillment. Many architects in the architectural industry, and even among the resource-based economy (RBE) movement, are more akin to "starchitects". The prefix "star-" is intended to mean that an architect has become a celebrity, a "star". Starchitects often copy protect their works. Most starchitects designs buildings as sculptures distinctive to their own subjective whims or the whims of a collective. Here, the term 'starchitect' is a pejorative. Besides copy protecting their work, starchitects create environments where people have to live in and among their artistic and subjective mental creations.

In modern commercialized society the architect designs the look, and the engineer determines how (if the design is novel) to make the look feasible. In community, the results of architecture are a whole integration into a materialized service system, wherein everything encoded into the system has been systematically engineered to do so. The distinction here is between starchitects in the market-State who (1) protect their efforts (thus, limiting cooperation and unification) and (2) do not account for others fulfillment in their work, and material design in a community-type society, where there is (1) no copy protection and (2) recognition of a social population with the potential for greater and lesser states of fulfillment, (and this, the effort to produce integrated and functional systems for that fulfillment).

INSIGHT: *In early 21st century society, architecture works on the principle of money: no money, no building; no money, no access to the architectural drawings.*

Here are some characteristics of the labor-market appearance of a starchitect:

1. Refusal to share their designs openly.
2. Disregard the effort and energy required to construct, maintain, disassemble, and clean the structure.
3. Disregard integration with the surrounding environment.
4. Disregard the integration of service functions, both within the structure itself and with concern to infrastructure.
5. Disregard nature-patterned aesthetics.
6. Disregard the nightmare/headache that an abstract piece of architecture will give to a structural engineer who has to design around the modernist structure.

The labor role of an architect is involves the building of structures for others. In community, we build structures for ourselves. Instead of a contextual world of harmonious geometric relationships and connectedness, architects tend to see a world of objects set apart from their contexts, with distinctive, attention-getting qualities. There are many such confirming studies. For example, Gifford et al. (2002) surveyed other research and noted that "architects did not merely disagree with laypersons about the aesthetic qualities of buildings, they were unable to predict how laypersons would assess buildings, even when they were explicitly asked to do so." The researchers traced this disagreement to well-known cognitive differences in the two populations: "Evidence that certain cognitive properties are related to building preference [was] found."

The division of construction by labor role has important consequences for the kinds of constructions that are produced into an environment. The same could be said for the field of engineering. Consider the engineering of living environments around the form of transportation we know as a "car", and the consequences that has had on our movement throughout life space engineered around vehicles. The architectural label, the architectural labor role, and socio-economic compartmentalization can prevent someone from seeing how certain designs disconnect and isolate people, and create hostile environments that cannot be shared well by people physically. Reward in service of effort is unhelpful; it decouples the designer from the context of human fulfillment.

An article by Mehaffy and Salingaros (2011) describes why, in the last half-century, the clear result of architectural construction is buildings whose makers have been so concerned with the drama of their appearance that they fail on the most fundamental human criteria. The following are some adapted quotes from the article.

> *Instead of a contextual world of harmonious geometric relationships and connectedness, architects (people providing the profession of architecture) tend to see a world of objects set apart from their contexts, with distinctive, attention-getting qualities. Their buildings celebrate the individuated form, as objects standing dramatically apart from context. Why do architects see the world in this unique way? Historically, and in part, this behavior is due to the conditioning present in architectural schooling (Gifford et al., 2002). Architectural students are typically asked to produce drawings that are pinned up next to one another, and then evaluated in a "crit" (or critique). In such an abstract setting, it is difficult for anyone to evaluate how well a project integrates with its context, if at all. Moreover, projects that are especially distinctive — object designs that stand out visually in an imaginative way by presenting an unusual structure — tend to get more attention from the faculty, and often, better grades. Those architects get rewarded, and selected out to be the later stars of the profession. Hence, architecture in the market-State has turned into something of a "novelty spectacle". However surprising and novel the forms of today's new architecture might appear, they remain tightly bound within this almost century-old model. The novelty spectacle approach has become the model not just for buildings, but also for whole cities.*
>
> *This focus on object-design has a deeper history in architecture. Up to about 1900, architects were understood to be practicing an adaptive craft, in which a building was an inseparable part of a dynamic living environment. "Blending in" respects the extant complex connective geometry, where components contribute to overall coherence. A building was assumed to meet the physiological and social needs of the people of that neighborhood first and foremost, and only then it would express its aesthetic qualities.*
>
> *Yes, our architect friends share much of the blame for the state of architecture, but let us remember that city officials, corporate executives, urban developers, mortgage bankers, and many others were part of this process of "architectural commodification", creating attention-getting product design rather than good sustainable environmental design. Clients, following what they took to be general consensus on what is great architecture, commissioned architects to build inhuman structures.*
>
> *All of these things are not, of course, trivial. They are the essence of a functional whole community, in which people are able to walk, navigate, feel well, and even feel any desire to live there in the first place. In short, the desires and gut reactions of the individual are the very essence of a great, living city, as opposed to a banal and dysfunctional one. The dysfunction of such image-based urban places — sadly all too common in the post-war era — is what has sent many people fleeing for the suburbs, with their simplistic ideas of retreat into a private garden. This too has turned into a dysfunctional failure of traffic congestion, blighted strip development, and isolated, car-dependent homes.*
>
> *In the early 21st century, many leading architects feel compelled to change the world drastically to make it conform to their preferred industrial paradigm. Unless non-architects (i.e., the rest of the population) stand up to this pressure, we risk the slow loss from attrition of all of humankind's most emotionally-nourishing creations. For example, architects see a well-functioning and beloved urban space but perceive it as ugly and offensive, desperately in need of immediate "re-qualification" to turn it into a contemporary hard industrial object. Politicians are happy to go along so as to please construction companies who profit from the unnecessary tearing down and rebuilding. The result is a sterile open space, unused, dysfunctional, and dead — but in the*

eyes of the architects, the operation has been a success!

Architects spend more time talking to their users, sharing their perception and understanding their needs: not just the architect's selfish need for artistic self-expression, or worse, his/her need to impress other architects and elite connoisseur-critics.

We are now dealing with an environment in which such image-based sculptural buildings are imposed upon people, whether they choose them or not. If such buildings "fall down on the job" of meeting human needs — if they are unduly stressful, or damaging to the quality of life — then that is a kind of architectural malpractice, and nothing less.

And thus we conclude that "architectural myopia" is a symptom of adopting a contradictory and opposite way of viewing the world. It also explains architects' insistence — continuous, strident, and bordering on the obsessive — of the need to "educate" the public. For every time public debate focuses upon the basic dichotomy in perceiving architectural form between architects and non-architects, the standard response by the former is to beg for more "education" of ordinary citizens, and to dismiss natural human responses to their work as being "unsophisticated" and "philistine". Architects really wish that normal people would undergo the same reversal, and then everyone might agree on the same non-contextual, non-adaptive building aesthetics.

Since the non-indoctrinated continue to see complexity and coherence in the living environment and refuse to accept "architectural myopia", the architect's strategy is simply to replace the built environment so that it no longer contains those essential elements of living structure.

For example, architects see a well-functioning and beloved urban space but perceive it as ugly and offensive, desperately in need of immediate "re-qualification" to turn it into a contemporary hard industrial object. Politicians are happy to go along so as to please construction companies who profit from the unnecessary tearing down and rebuilding. The result is a sterile open space, unused, dysfunctional, and dead — but in the eyes of the architects, the operation has been a success!

A culture based upon an abstract, disconnected conception of space is re-shaping our world right now for the worse. The parallel reality is replacing the living one. Enthusiastically supported by politicians and the building industry, architects have been commissioned to destroy historic buildings and urban spaces worldwide. Because "architectural myopia" is justified as perfectly normal in the press, such interventions are praised by their promoters but turn out disastrous for the urban fabric, and are hated by potential users. Those projects all tend to look and feel the same. This is not surprising, since the designs are generated by the same abstract modernist images in the minds of architects oblivious of the connective geometry that would catalyze the eventual life in such a space.

We desperately need a new kind of architect: one more focused on process than on systems and context, rather than just objects. We need an architecture that actually optimizes the user (human) experience of the built environment.

Firstly, it is important to re-integrate the needs of human beings, including their sensory experience of the world, and their participation into the process of designing buildings and cities. Preparing our new type of architect for practice, we should re-examine the ways that architects are rewarded today: the corrupt and incestuous system of financial incentives, corporate branding, and image-making that rewards the extravagant "starchitect" over the contextual practitioner. Once we have created a consensus for radical change, it will be straightforward to find new ways of compensating good work, through more incentives such as awards, commissions, scientific research that identifies both successes and failures, and other, stronger feedback.

Scholarly references

- Gifford, R., Hine, D.W., Muller-Clemm, W., Shaw, K.T., (2002). *Why architects and laypersons judge buildings differently: cognitive properties and physical bases.* Journal of Architectural and Planning Research, 19(2), pp131-148. [jstor.org]
- Salingaros, N.A. *Nikos A. Salingaros: articles on architecture, complexity, patterns, and urbanism.* Accessed: January 7, 2020. [zeta.math.utsa.edu]

Online references

- Mehaffy, M., Salingaros, N.A. (2011). *The Architect Has No Clothes: Why so much modern design looks harsh and feels inhospitable*. On The Commons. [onthecommons.org]

Life Support: Water Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: Water, hydro, hydro service, water service, water system, hydrological system

Abstract

Water is called the "universal solvent" because it is capable of dissolving more substances than any other liquid. This is important to every living thing on earth. Water is a core component of the planet's ecospheric cycle. Hydrology is the study of water in its totality, including is cycles and qualities. Water is an essential requirement of all life. Water is a uniquely accountable physical resource (Read: object). Different needs, and sub-services therein, have different requirements of and for water. Water is a material with many different sub-compositions, some of which are usable for specific functions, and others of which are not. Water serves many functions in a materialized societal system. The the processing and distribution of water can be integrated into a city platform.

Graphical Abstract

No significant image available

1 Hydrology and the earth's hydrological/water cycle

Hydrology is the scientific study of the movement, distribution, and quality of water on Earth and other planets, including the hydrologic cycle, water resources and environmental watershed sustainability. Hydrology is subdivided into surface water hydrology, groundwater hydrology (hydrogeology), and marine hydrology. Domains of hydrology include hydrometeorology, surface hydrology, hydrogeology, drainage basin management and water quality, where water plays the central role. In terms of environmental measurements, hydrology refers to the physical movement of a body of water, including changes in water level, flow, and other dynamic processes. Hydrological modeling refers to modeling of the hydrologic cycle (i.e., water cycle). The water cycle, also known as the hydrologic cycle or the H2O cycle, describes the continuous movement of water on, above and below the surface of the Earth. The hydrologic cycle is the term used to describe the natural circulation of "raw" water in, on, and above the earth. Water occurs in many forms as it moves through this cycle. The Earth's water cycle is sub-composed of a set of identifiable processes/steps listed in table ...

Water is placed in the air by evaporation from water and land surfaces and by transpiration from plants. It then condenses to produce cloud formations and returns to earth as rain, snow, sleet, or hail. Some of this evaporates, while some flows as runoff into lakes and streams. The remainder goes into the soil and then into underlying rock formations by seepage or infiltration. The water which has seeped through the earth will finally find its way to the surface through springs. It can also flow through porous media until intercepted by streams, lakes, or oceans. The cycle does not always progress through a regular sequence; steps may be omitted or repeated at any point. For example, precipitation in hot climate may be almost wholly evaporated and returned to the atmosphere.

> **NOTE:** *With the appropriate technology, water may be collected at any point in the Earth's water/hydrological cycle.*

The mass of water on Earth remains fairly constant over time but the partitioning of the water into the major reservoirs of ice, fresh water, saline water and atmospheric water is variable depending on a wide range of climatic variables. The water moves from one reservoir to another, such as from river to ocean, or from the ocean to the atmosphere, by the physical processes of evaporation, condensation, precipitation, infiltration, runoff, and subsurface flow. In doing so, the water goes through different phases: liquid, solid (ice), and gas (vapor). The water cycle involves the exchange of energy, which leads to temperature changes. For instance, when water evaporates, it takes up energy from its surroundings and cools the environment. When it condenses, it releases energy and warms the environment. These heat exchanges influence climate.

1.1 *Primary components of the water service system*

The water system is a system of engineered hydrologic and hydraulic components that provide a water supply, modified where necessary, to meet service requirements.

A water supply system involves everything from the collection of source water through to the point of use, as well as all transformation, recycling, and transportation therein.

A water supply system typically includes:

1. An **ecosystem** that re-cycles water and fulfills the H20 needs of organisms living in a habitat. The ecosystem of a water body is known as an 'aquatic ecosystem'.
2. A **drainage basin** (**catchment basin)** is an extant or area of land where surface water from rain, melting snow, or ice converges to a single point at a lower elevation, usually the exist of the basin where water joins another water body, such as a river, lake, reservoir, estuary, wetland, sea, or ocean. Other terms that are used to describe drainage basins are catchment, catchment area, drainage area, river basin and water basin. The drainage basin includes both the streams and rivers that convey the water as well as the land surfaces from which water drains into those channels, and is separated from adjacent basins by a drainage divide. The drainage basin acts like a funnel, collecting all the water within the area covered by the basin and channelling it into a waterway. Each drainage basin is separated topographically from adjacent basins by a geographical barrier such as a ridge, hill or mountain, which is known as a water divide. In nature, drainage basins are dynamic because the water and sediment-producing areas expand and contract depending upon the catchment characteristics, the antecedent conditions prior to any water even, and the character of the water input. In the technical sense, a **watershed** refers to a divide that separates one drainage area from another drainage area. However, in some areas, the term is often used to mean the drainage basin or catchment area itself.
3. A **raw water collection point** (above or below ground) where the water accumulates, such as a lake, a river, or groundwater from an underground aquifer. Water in this form is considered "raw", as opposed to water which has been treated before consumption, such as "drinking water" or

water which has been used in an materialization process, such as "waste water". Raw water may be transferred using uncovered ground-level aqueducts, covered tunnels or underground water pipes.

4. **Water processing and purification** technologies.
5. **Water storage** locations such as reservoirs, water tanks, or water towers. Smaller water systems may store the water in cisterns or pressure vessels. Tall buildings may also need to store water locally in pressure vessels in order for the water to reach the upper floors.
6. Additional **water pressurizing components** [beyond that of gravity due to elevation differentials] such as pumping stations, which may need to be situated at the outlet of underground or above ground reservoirs or cisterns (if gravity flow is impractical).
7. A **pipe network for distribution/transportation** of water to points of storage, transformation, or usage. Water is transferred using water pipes (usually underground).
8. **Input and output connections** the distribution/ transportation network, including drains and faucets.

1.2 Raw water sources and catchment

The collection of water from a raw water source is referred to as **raw water collection/catchment**. The **Earth's water cycle** provides sources of water. Those sources of water are:

1. **Ground sources:** include groundwater, springs, hyporheic zones and aquifers. **Ground water** is water that moves in some manner through the "ground" (i.e., water that has "gone underground"). **Spring water** is water that has moved through the ground and reappeared on the surface of the land. If you travelled underground you would eventually get deep enough to find that all the rock around you is soaked with water. You'd have entered the "saturated zone". The height of water in the saturated zone is called the **water table**. In dry places, the water table is very deep, but in moist places, the water table is very shallow. When the water table is higher than the actual surface of the ground, there are streams, rivers, and lakes on the land (i.e., surface water). Ground water is partially "filtered" through the ground, and partially flows through rivers that disappear beneath the earth. Rain that soaks into the ground, rivers that disappear beneath the earth, melting snow are but a few of the sources that recharge the supply of underground water. Because of the many sources of recharge, ground water may contain any or all of the contaminants found in surface water as well as the dissolved minerals it picks up during its long stay underground. Collection technologies: welling technologies and distribution technologies (e.g., applied hydraulics and pumps).
2. **Surface water:** include rivers, streams, glaciers, including some a superficial cavity of some kind, like a dug well. As water travels over the surface of the land, it dissolves naturally occurring minerals and can pick up substances resulting from the presence of animals or from human activity. Surface water is exposed to many different "contaminants", such as foliage, wastes, -icides, algae, and many other organic and inorganic materials. Collection technologies: land restructuring technologies; catchment technologies; evaporation technologies; distribution technologies
3. **Precipitation:** includes rain, hail, snow, fog, etc. Precipitation is any form of liquid or solid water particles that fall from the atmosphere and reach the surface of the Earth. Precipitation comes in many forms, but it all comes from the same general process. Collection technologies: rain water catchment technologies (e.g., rain barrels) and distribution technologies. Any surface with significant displacement is probably a good water catchment platform. Roofs and trees, for example, work for water catchment. Water rains down on the roof and it is channelled into a gutter that feeds into the water system.
4. **Biological sources:** such as plants. Some plants directly produce a drinking "water" liquid. One of the most well known of these plants is the coconut palm tree. However, drinking water can be collected from the transpiration of water. Most plants constantly transpire water vapor. The transpiration process leads to the release of water vapor from plants and soil into the air. The intent of this technique is to collect and condense plant-respirated water vapor. This is often done by placing a container over the plant and securing the opening. Collection technologies: desalination technology; evaporation technology; distillation technology; osmotic technology.

- **Ocean water (salinated seawater):** - About 97 percent of the Earth's water can be found in the unified form of an ocean. The ocean covers about 71 percent of the Earth and is blue, while land makes up the other 29 percent and varies in colour. The large presence of an ocean gives the Earth the appearance of a blue marble. Collection technologies: desalination technology; evaporation technology; distillation technology; osmotic technology. Ocean breezes contain elements which are highly corrosive to various substances.

5. **Atmospheric sources:** Water is constantly evaporating into the air. Air can carry water, it is called 'humidity'. All breathable air has some degree of water vapor (moisture) present.
   A. **Relative humidity (RH)** is the amount of water vapor in the air. Water vapor is the gaseous state of water and is invisible to the human eye. Humidity indicates the likelihood of precipitation, dew, or fog. While humidity itself is a climate variable, it also interacts strongly with other climate variables. Humidity depends on water vaporization and condensation, which, in turn, mainly depends on temperature. Humidity is affected by winds and by rainfall. The most humid regions of the earth are generally located closer to the equator, near coastal regions. Relative humidity is not a good measure for how humid it feels outside. Humid air is less dense than dry air. Dewpoint is a more accurate measure of how humid it feels.
   B. **Dewpoint** is a measure of atmospheric moisture. The higher the dewpoint, the more moisture in the air. It is the temperature measure to which air must be cooled to reach saturation (assuming air pressure and moisture content are constant). A higher dew point indicates more moisture present in the air. It is sometimes referred to as dew point temperature, and sometimes written as one word (dewpoint). Frost point is the dew point when temperatures are below freezing.

1.3 Functional usages of water

The parameters for water quality are determined by the intended use. Work in the area of water quality tends to be focused on water that is treated for human consumption, industrial use, or in the environment.

1. Drinking water (a.k.a., potable water/"spring water") - this is water that is safe for humans to drink and to use in nutrition preparation (i.e., food preparation).
2. Ecological cultivation and aquaculture - water for consumption, or a life-media, by other organisms in the ecology.
3. Good and service production (including medicine, cleaning, and other productive uses)
4. Washing & cleaning.
5. Recreating in (i.e., swimming ponds and pools).
6. As a "waste" recycling medium.
7. Hydronic radiant heat to flooring.
8. Power/Energy production - water turbine and electrolysis (water splitting) to produce hydrogen.
9. Power/Energy emergency control - fire suppressant.

NOTE: *Electrolysis is the process of passing electricity through water and splitting the water into its component parts.*

1.3.1 Hydraulic water distribution, water movement & distribution

Water is moved around a water system through either pumping or gravity-flow through elevation differentials. Therein, elevation can be used to run distribution sub-systems through gravity pressure.

1. Positive pressure generated by gravity through a water tower. Great big tanks of water up on stilts. Pumps run all day to push water up into them, and when you open the faucet water flows downhill through the pipes and out of the tap. A tall tank is necessary to provide pressure, but it doesn't have to be big to provide quantity, as the water may be pumped up just before use.
2. Gravity flow and pump flow. A water service system could use one or the other, but generally they are used together.
3. There are other ways of providing pressure than gravity flow. A pump can provide positive pressure. Or a pump filling an internal airspace at the top of a tank could provide positive pressure. Pumps are located in "pump stations".
4. Reservoirs or when the water source is at a higher elevation create an environment where water flows downhill by gravity, through processing, into the mains distribution, with no pumps required. In some higher elevation geographic locations the water must be pumped up to the higher elevation to maintain enough pressure.
5. If there is too much pressure due to elevation, then sometimes the water processing system downstream has to decrease pressure rather than increasing it.
6. Pumps and water towers work hand-in-hand. Without water towers, the pumps would have to operate continuously, even during low-demand periods. With water towers, the pumps pump water up into the towers, which then maintain pressure. As water is consumed, the level in the water tower drops until it reaches a certain point, which actuates a level switch to start up the pumps again. You can't have water towers without pumps. You can have pumps without water towers, but would always need to keep the pumps running (which is not generally practical).

7. Today speed drives control the speed of the pumps. As demand goes up the pumps speed up. Depending on the city the system pressure can run from 50 to 120 PSI.
8. Sometime buildings in a water distribution system have water regulators in them to protect the pipes of the building from the city's higher pressure.
9. However, it's pretty easy to provide the same constant, even, reliable pressure if you let gravity do it, or enclose the water in a tank which also contains a compressed gas like air -- a substance which is highly compressible.
10. You pump water into a tank, even in spurts, or compress air into a tank at irregular intervals, and the resulting gravity or air pressure evens out the flow to a workable level (e.g., positive-displacement pump).
11. When you have a centralized network around a series of main pumps, then the pressure will likely taper off as you approached the periphery of the network. Hence, elevated storage tanks and re-pump stations are necessary at strategic locations.
12. A clear well (a.k.a., clearwell) is an output/finished water storage area.

1.4 Water storage points

Water can be stored in the following ways:

- Underground reservoirs; lakes; ponds; streams
- Water channels, wells, and pools
- Water containers including barrels, tanks, and other vessels
- Water pipes
- Water bottles

1.5 Water distribution access points

Water access points are the locations where water is accessible to systems, to the community, and to individuals.

Common water distribution access points include, but are not limited:

1. Buildings where water is used.
2. Portable stores of water (i.e., bottled water).
3. Swimming points.
4. Places where people move a lot for hydration and recreation (e.g., bodies of water and water fountains).
5. Ecosystem water distribution points (i.e., sprinklers).
6. Fire suppressant system distribution points.

1.6 Human-use water outlets (taps & faucets)

Human use water switches are designed to be as precise, simple, and helpful as possible. Faucets have three primary functional designs:

1. No handle design (i.e., touch interface or other).
2. Single handle design allow the user to turn on and off the tap with one hand. Users of single handle designs are less likely to leave the water on than users of a two handle design. Also, with a single handle design the user can set the temperature of the water more easily than with two handles.

1.7 Water as a solvent

Water is one of the most important elements humankind needs to survive. It is a universal solvent for living things, which can dissolve many substances, both solid and gaseous, that come into contact with it. We use water everywhere, constantly.

The temperature of water affects how quickly substances dissolve in it and, in some cases, the quantity that can be dissolved: for example, warm water holds less dissolved oxygen than colder water, but it holds more of most solids. Temperature also has a major effect on living things within the water. Compared to air or soil, bodies of water change temperature slowly, so aquatic life is generally not exposed to sudden fluctuations of temperature.

2 Water quality

Water quality testing is an important part of environmental monitoring. When water quality is poor, it affects not only aquatic life but the surrounding ecosystem as well. By measuring the characteristics of water bodies we can determine its ability to sustain life and meet our productive and ecological needs.

"Pure" water is essentially non-existent in the natural environment. Natural water, whether in the atmosphere, on the ground surface, or under the ground, always contains dissolved minerals and gases as a result of its interaction with the atmosphere, minerals in rocks, organic matter, and living organisms. Water is dangerously good at dissolving things. Since water is a polar molecule, its positive end is attracted to negatively charged ions or the negative sides of other polar molecules, and its negative side is attracted to positively charged ions or the positive sides of other polar molecules.

Impurities in raw water are either **suspended** or **dissolved**. Suspended impurities include diseases organisms, silt, bacteria, and algae. Some disease organisms must be removed or destroyed before the water is safe to consume. Dissolved impurities include salts, (calcium, magnesium, and sodium), iron, manganese, and gases (oxygen, carbon dioxide, hydrogen sulfide, and nitrogen). These impurities must be reduced to levels acceptable for human consumption.

As water goes through the hydrologic cycle, it gathers many impurities. Dust, smoke, and gases fill the air and can contaminate rain, snow, hail, and sleet. As runoff, water picks up silt, chemicals, and disease organisms. Water is a carrier of many organisms which cause disease. As it enters the earth through infiltration, some of the suspended impurities may be filtered out. However, other minerals and chemicals are dissolved and carried along. As ground water, in underground reservoirs, it may contain disease organisms as well as harmful chemicals. In addition to the impurities in water resulting from infiltration, many are contributed by an industrialized society. Garbage, sewage, industrial waste, pesticides, and nuclear, biological, and chemical (NBC) agents are all possible contaminants of raw water.

> **DESIGN REQUIREMENT:** *Drinking water must be free of anything that would degrade the human restoration and performance cycle. Additionally, it should not damage the materials used in its transportation and storage.*

In community, we scientifically monitor all water sources for quality (i.e., composition and condition). Generally, ground water (subsurface) and spring water has less chemical and biological contaminants than surface water, provided reasonable care is exercised in the selection of the well site or spring. Harmful microorganisms are usually reduced to tolerable levels by passage through the soil; however, the water should still be tested. Water collected from its raw source does not always require purification/modification.

Water quality is a measure of the condition of water relative to the requirements of one or more biotic species and or to any human need or purpose. It is most frequently used by reference to a set of standards. The parameters for water quality are determined by the intended use.

Environmental indicators and parameters for water quality typically fall under six categories:

1. **Physical properties/indicators**
    - Temperature, specific conductance (EC), electrical conductance, conductivity, total dissolved solids (TDS) + hardness, total suspended solids (TSS), turbidity (a.k.a., transparency), odour, colour, taste, volume and depth
- **Chemical composition/indicators**
    - Potential hydrogen (pH), biochemical oxygen demand (BOD), chemical oxygen demand (COD), dissolved oxygen (DO) + oxidation-reduction potential (ORP), total hardness (TH), metals and metalloids, hormone analogues, pharmaceuticals, metabolites, surfactants, and all other chemical compounds organic/in-organic/ natural/man-made which may be found in the water
2. **Biological composition/indicators** (i.e., ecological assemblage - the collection of organisms in the aquatic ecology)
    - Algae, phytoplankton, vegetation, fish, reptiles, insects, and all other aquatic life
3. **Microbiological composition/Indicators**
    - Bacteria, protozoa, viruses
4. **Radiological composition/indicators**
    - Radon, nuclear waste, nuclear medicines, and all other radioactive elements
5. **Quantum electrodynamic coherency**
    - Structure + temperature of water molecules
6. **Movement of water**
    - The properties of water vary naturally depending on the surrounding environment; hence, data collected about water quality must be interpreted in the context of the water body's particular environment and position in the catchment.
    - **Flow** is the volume of fluid (in this case, water) that passes through a passage of any given section in a unit of time. Flow, for the purposes of measurement, is the **velocity** of water multiplied by the cross-sectional area of the stream. Flow is modified by conditions along and around a waterway. The amount of any particular substance carried in the water is known as the **load**.

Biological assessment (bioassessment) is an evaluation of the condition of a water body based on the organisms living within it. It reveals cumulative effects, as opposed to chemical observation, which is representative only at the actual time of sampling. Biological systems reflect overall ecological integrity and integrate the effects of different stressors. Thus, their observation provides a broad measure of aggregate impact and fluctuating environmental conditions. The primary reason for bioassessment and monitoring is that degradation of water body habitats affects the biota using those habitats; therefore, the living organisms themselves provide the most direct means of assessing real environmental impacts.

2.1 Definition of indicators

Physical indicators include:

1. **Temperature** - the degree or intensity of heat present in a substance or object, especially as expressed according to a comparative scale and shown by a thermometer or perceived by touch. Temperature is a physical property of water that quantitatively expresses hot and cold. The temperature of a water body directly affects many physical, biological and chemical characteristics. Temperature is measured in Celsius.
    - Water changes phase at specific temperatures (accounting for pressure).
    - The temperature of a body of water influences its overall quality and ecology. Warm waters are more susceptible to eutrophication — a build-up of nutrients and possible algal blooms — because photosynthesis and bacterial decomposition both work faster at higher temperatures. Oxygen is less soluble in warmer water and this can affect aquatic life. By contrast, salts are more soluble in warmer water, so temperature can affect the water's salinity. Temperature directly affects the metabolic rate of plants and animals. Aquatic species have evolved to live in water of specific temperatures. If the water becomes colder or warmer, the organisms do not function as effectively, and become more susceptible to toxic wastes, parasites and diseases. With extreme temperature change, many organisms will die. Changes in long-term temperature average may cause differences in the species that are present in the ecosystem.
    - If the water temperature changes by even a few degrees, it could indicate a source of unnatural warming of the water or thermal pollution.
    - Factors that affect water temperature: air temperature; air movement; amount of shade and sunlight; soil erosion increasing turbidity; thermal pollution; water depth and volume; confluence of water movement.
    - Effects of water temperature: suspended and dissolved solids are changed; solubility of dissolved gas; rate of plant growth; metabolic rate of organisms; resistance of organisms.
2. **Electrical conductivity (EC)** is the property of a substance which enables it to serve as a channel for medium for electricity. It is a measure of water's capability to pass electrical flow. This ability is directly related to the concentration of ions in the water. These conductive ions come from dissolved salts and inorganic materials such as alkalis, chlorides, sulfides and sulphates, and carbonate compounds. Compounds that dissolve into ions are also known as electrolytes. The more ions that are present, the higher the conductivity of water. Likewise, the fewer ions that are in the water, the less conductive it is. Distilled or deionized water can act as an insulator due to its very low (if not negligible) conductivity value. Sea water, on the other hand, has a very high conductivity. Conductivity, in particular specific conductance, is one of the most useful and commonly measured water quality parameters. In addition to being the basis of most salinity and total dissolved solids calculations, conductivity is an early indicator of change in a water system. Salty water conducts electricity more readily than purer water. There are a number of scales used in EC, most commonly micro-Siemens (µS) or milli-Siemens (mS). For example, if a particular application calls for water with "2.0 EC," this is an incorrect determination. Most likely, the application is calling for an EC level of 2.0 mS. 2.0 mS = 2000 µS.
3. **Total Dissolved Solids (TDS)** is a measure of the combined/total amount of mobile charged ions, including minerals, salts or metals dissolved in a given volume of water, expressed in units of mg per unit volume of water (mg/L), also referred to as parts per million (ppm). TDS is a measure of the combined content of all inorganic and organic substances. The TDS of water is composed of mineral salts and small amounts of other inorganic and organic substances contained in a liquid in molecular, ionized or micro-granular (colloidal sol) suspended form. A TDS measure combines the sum of all ion particles that are smaller than 2 microns (0.0002 cm). Generally the operational definition is that the solids must be small enough to survive filtration through a filter with two-micrometer (nominal size, or smaller) pores. There are three primary classifications of water, based on the concentration of total dissolved solids:

- **Fresh** - TDS less than 1,000 - 1,500 ppm (milligrams-per-liter of dissolved solids).
- **Brackish** - TDS between 1,500 ppm and 16,00 ppm; Brackish water (less commonly brack water) is salt water and fresh water mixed together.
- **Salt water** (seawater) - TDS greater than 15,000 ppm.

4. **Hardness** in water is caused by calcium (Ca^{2+}) and magnesium (Mg^{2+}) - two nontoxic, naturally occurring minerals in water. Water hardness is typically reported in grains per gallon, milligrams per liter (mg/l) or parts per million (ppm). One grain of hardness equals approximately 17.1 ppm (mg/L) in TDS. Excessive hardness makes it difficult for soap to lather, leaves spots on dishware, reduces water flow, and can cause pipe, valve, and drain "scaling".
   - **Soft water** is water with a low mineral content. **Water softening** is the removal of calcium, magnesium, and certain other metal cations in hard water. The resulting soft water is more compatible with soap and extends the lifetime of plumbing.
   - **Hard water** is water with a high mineral content.
5. **Turbidity** is the measure of relative clarity of a liquid. It refers to the opacity or muddiness caused by particles of extraneous matter. Turbidity is measured in nephelometric turbidity units (NTU). The instrument used for measuring it is called nephelometer or turbidimeter, which measures the intensity of light scattered at 90 degrees as a beam of light passes through a water sample. In general, the more material that is suspended in water, the greater is the water's turbidity and the lower its clarity. Turbidity affects how far light can penetrate into the water. It is not related to water colour: tannin-rich waters that flow through peaty areas are highly coloured but are usually clear, with very low turbidity. Measures of turbidity are not measures of the concentration, type or size of particles present, though turbidity is often used as an indicator of the total amount of material suspended in the water (called total suspended solids). Turbidity can indicate the presence of sediment that has run off from construction, agricultural practices, logging or industrial discharges.
   - Suspended particles absorb heat, so water temperature rises faster in turbid water than it does in clear water. Then, since warm water holds less dissolved oxygen than cold water, the concentration of dissolved oxygen decreases.
   - If penetration of light into the water is restricted, photosynthesis of green plants in the water is also restricted. This means less food and oxygen is available for aquatic animals. Plants that can either photosynthesise in low light or control their position in the water, such as blue-green algae, have an advantage in highly turbid waters.
   - Regular turbidity monitoring may detect changes to erosion patterns in the catchment over time. Event monitoring (before, during and immediately after rain) above and below suspected sources of sediment can indicate the extent of particular runoff problems.

Chemical indicators include:

1. **pH (potential hydrogen) -** The measured indicator for acidity or alkalinity is known as the pH value. A pH value of 7 means a substance is neutral, water with a pH lower than 7 is considered acidic and with a pH greater than 7 water is considered alkaline. The normal range for pH in surface water systems is 6.5 to 8.5. Potential hydrogen is measured with a pH meter.
2. **Dissolved oxygen (DO)** refers to the level of free, non-compound oxygen present in water or other liquids. It is an important parameter in assessing water quality because of its influence on the organisms living within a body of water. In limnology (the study of lakes), dissolved oxygen is an essential factor second only to water itself. A dissolved oxygen level that is too high or too low can harm aquatic life and affect water quality. Dissolved oxygen is a measure of the quantity of oxygen present in water (it has nothing to do with the oxygen atoms within the water molecules).
   - Oxygen is essential for almost all forms of life. Aquatic animals, plants and most bacteria need it for respiration (getting energy from food), as well as for some chemical reactions. The concentration of dissolved oxygen is an important indicator of the health of the aquatic ecosystem. Persistently low dissolved oxygen will harm most aquatic life because there will not be enough for them to use.
   - In some circumstances, water can contain too much oxygen and is said to be supersaturated with oxygen. This can be dangerous for fish. Supersaturated conditions occur in highly turbulent waters in turbines and at spillways, because of aeration, and also on sunny days in waters experiencing algal blooms or with many aquatic plants, because of photosynthesis. In this supersaturated environment, the oxygen concentration in fishes' blood rises. When the fish swim out into water that has less dissolved oxygen, bubbles of oxygen quickly form in their

blood, harming the circulation.
- The air is one source of dissolved oxygen, and aquatic plants, including algae, are another. The speed at which oxygen from the air enters and mixes through a water body depends on the amount of agitation at the water surface, the depth of the water body and the rate at which it mixes itself. As water temperature rises, oxygen diffuses out of the water into the atmosphere.
- Dissolved oxygen concentrations change with the seasons, as well as daily, as the temperature of the water changes. At very high altitudes, the low atmospheric pressure means dissolved oxygen concentrations are lower
- Deep muddy lowland rivers, which contain more organic matter than upland streams, are likely to have lower dissolved oxygen concentrations than upland streams because bacteria are using the oxygen to break down the organic matter. Likewise, dissolved oxygen is usually lower than normal after storms have washed organic materials into any water body. Aquatic plants photosynthesise during daylight and increase dissolved oxygen concentrations around them.

3. **Oxidation-reduction potential (ORP)** is a measurement of water's ability to oxidize contaminants. The higher the ORP, the greater the number of oxidizing agents. Checking ORP is a simple method to monitor the effectiveness of a sanitizer or the quantity of anti-oxidants in a liquid. The ORP level of water can also be viewed as the level of bacterial activity of the water because a direct link occurs between ORP level and the count of certain bacterial species in water. ORP sensors work by measuring the dissolved oxygen. An ORP meter measures very small voltages generated with a probe placed in water. The electrode is made of platinum or gold, which reversibly looses its electrons to the oxidizer. A voltage is generated which is compared to a silver (reference) electrode in a silver salt solution, similar to a pH probe. The more oxidizer available, the greater the voltage difference between the solutions. ORP cannot be used as a direct indicator of dissolved ozone residual, except in very clean water applications. If there is ozone in the water, then ORP is a convenient measure of ozone's ability to perform a chemical task.

2.2 Water quality standards

Drinking water standard:

- Drinking-water should be supplied under continuous positive pressure in a plumbing system free of any defects that could lead to contamination of any product.
- Drinking-water is unmodified.
- Water that has elements or compounds with additive pharmacological effects (e.g., sodium fluoride/fluoridation) is not pure drinking water.

Radiological standards for water include:

- Radiological water quality standards are based on the fact that radiation has an adverse physical effect on human physiology. Any treated water that contains nuclear contamination should be avoided. When ingested, radioactive isotopes interfere with the reproduction of human cells. They can cause nausea, vomiting, and hair loss and weaken the body's defenses to infections.

3 Water processing

NOTE: *Nature makes water "pure" by heating it up, causing evaporation, then eventually condensing into clouds, and finally, falling back to the earth as rain (or snow). However, when water passes through the atmosphere it will pickup pollutants present therein.*

The objective of water processing is to produce water that meets the specified requirements of a specific purpose. Water may be processed so that it is fit for human consumption (as "drinking water"). Water processing may also be designed for a variety of other purposes, including fulfilling the requirements of medical, chemical, and other applications. The most well-known water process is that of water filtration. However, water processing may also include the restructuring of water and the addition of elements to the water, as well as modification of the water's properties, including pH and temperature.

In industry, the processes that make water more acceptable for an end-use, which may be drinking, industry, or medicine are called "water treatment". The term "water treatment" generally refers to potable water production from raw water, whereas "wastewater treatment" refers to the treatment of polluted water, where the pollution could be from human waste, industry, agricultural waste or other sources of pollution.

Many people assume that the water flushed down their drains is of little consequence to the environment, with any contaminants of concern being removed by wastewater treatment plants. In reality, however, most wastewater treatment plants are not equipped to remove various medications and other chemicals that may end up being flushed or poured down your drain. They end up being continuously recycled through a populations water supply and accumulate over time.

There are many different "grades" of water. The common classifications for water are:

1. **Drinking water (drinking-water)** is water that is suitable for drinking. Drinking water is both potable and palatable. Drinking water is the minimum quality of water that should be used for the preparation of productive substances. Once produced, drinking water has to be protected from microorganisms and contaminants that can enter the water through the pipelines that transport it to households. This happens by means of disinfection with various disinfection agents.
   A. **Potable water** (meaning drinkable or fit to drink) is water that is free from disease-producing organisms, poisonous and toxic substances, chemical or biological agents, and radioactive elements, which make it unfit for human consumption or other uses.
   B. **Palatable water** is water that is pleasing in appearance and taste. It is significantly free from colour, turbidity, taste, and odor. It should also be cool and aerated. Water may be palatable and at the same time not be potable. However, visual inspection cannot determine if water is of appropriate quality to drink.
2. **Filtered water** is water that has been processed at a basic level to remove chemicals, biologicals, suspended solids, and gases.
3. **Purified water** is water that has been processed to meet the requirements for ionic and organic chemical purity, and is protected from microbial contamination. In other words, purified water is generally free of dissolved minerals. While not all filtered water is purified water, all purified water has been filtered. The minimum quality of source or "feed water" for the production of purified water is drinking water. "Purified" water, being essentially free of minerals, is an active absorber (i.e., it is "aggressive"), in that it tends to dissolve substances with which it is in contact. Notably, when it comes into contact with air, carbon dioxide from the air is rapidly absorbed, making the water acidic and even more "aggressive". Many metals are dissolved by purified water.
4. **Black water (blackwater)** is water that contains animal, human, or food waste. Black water is also referred to as sewage, wastewater and sewer water. Blackwater generally travels from where it is produced as sewage to either an on-site septic system or a water processing system prior to its release into an ecosystem.
5. **Grey water (greywater)** is used water from bathroom sinks, showers, tubs, and washing machines. It is not water that has come into contact with faeces, either from the toilet or from washing diapers. Grey water may contain traces of dirt, food, grease, hair, and biodegradable cleaning products.
6. **Waste water (wastewater)** is any water that has been adversely affected in quality by anthropogenic influence.
7. **Recycled/reclaimed water** is water that has been processed and captured for reuse.

Water system processing service points:

1. Source-point
2. Continuous storage
3. End-point
4. Mobile/portable

3.1 Water Purification processes

Typical water purification processes require:

1. **Aeration** is the process of increasing the oxygen saturation of the water. Dissolved oxygen (DO) is a major contributor to water quality and a factor in the pH of water. Not only do fish and other aquatic animals need it, but oxygen breathing aerobic bacteria decompose organic matter. When oxygen concentrations become low, anoxic conditions may develop which can decrease the ability of the water body to support life. Water aeration is often required in water bodies that suffer from anoxic conditions, usually caused by adjacent human activities such as sewage discharges, agricultural run-off, or over-baiting a fishing lake. Aeration is one method of for reducing algae growth in a water body. Any procedure by which oxygen is added to water can be considered a type of water aeration. This being the only criterion, there are a variety of ways to aerate water. These fall into two broad areas – surface aeration and subsurface aeration. There are a number of techniques and technologies available for both approaches. Natural aeration is a type of both sub-surface and surface aeration. It can occur through sub-surface aquatic plants. Through the natural process of photosynthesis, water plants release oxygen into the water providing it with the oxygen necessary for fish to live and aerobic bacteria to break down excess nutrients. Oxygen can be driven into the water when the wind disturbs the surface of the water body and natural aeration can occur through a movement of water caused by an incoming stream, waterfall, or even a strong flood. In large water bodies, autumn turn-over can introduce oxygen rich water into the oxygen poor hypolimnion.
   - A solar aeration unit (i.e., solar aerator) is a pump powered by solar energy that aerates a water body. Fountains are a form of aeration.
2. **pH adjustment** is typically done with chemicals, such as caustic soda, lime, soda ash, or sodium bicarbonate. One natural way to raise the pH of the ground water is through aeration. Aeration will strip off dissolved carbon dioxide and raise the pH of the water. If there is radon present in the water, it will also be stripped.
3. **Ion exchange (IE or ix)** is a chemical water processing method involving reversible chemical reactions for removing dissolved ions from solution and replacing them with other similarly charged ions. In other words, undesirable contaminants are removed from water by exchange with another non-objectionable, or less objectionable substance. Both the contaminant and the exchanged substance must be dissolved and have the same type (+,-) of electrical charge. **Deionized water** (DI water, DIW or de-ionized water), often confused with demineralized water / DM water, is water that has had almost all of its mineral ions removed, such as cations like sodium, calcium, iron, and copper, and anions such as chloride and sulfate. Because most non-particulate water impurities are dissolved salts, **deionization** produces a high purity water that is generally similar to distilled water, and this process is quick and without scale buildup. However, deionization does not significantly remove uncharged organic molecules, viruses or bacteria, except by incidental trapping in the resin. Deionization can be done continuously and inexpensively using **electrodeionization**. Three types of deionization exist: co-current, counter-current, and mixed bed. **Water softening** is the removal of calcium, magnesium, and certain other metal cations in hard water. The resulting soft water is more compatible with soap and extends the lifetime of plumbing. Water softening is usually achieved using lime softening or ion-exchange resins. As water flow through the water softener, it will pass through a resin, bed of small plastic beads or chemical matrix (called Zeolite) that will exchange the calcium and magnesium ions with sodium ions (salt). Therefore, the TDS level will remain virtually constant (there may be minor differences).
4. **Dilution** is a natural process that occurs in nature when "pollution" occurs. Nature provides its flowing waters with the ability to restore themselves through their own self-purification processes. It was only when humans gathered in great numbers to form cities and created novel chemicals that the stream systems were not always able to recover from having received great quantities of waste. Note, the self-purification discussed here relates to the purification of organic matter only.

3.1.1 Emergency water purification

Each individual should select the clearest, cleanest water with the least odor and then process the water using individual water purification procedures from their knowledge skill-set. Productive technologies for emergency water purification include: bottled water; iodine tablets, chlorine ampules, boiling, earth-based filtration, hand operated water purifiers; spring water; and passive water generation techniques.

3.2 Water processing technologies

Disinfection and sterilization are both decontamination processes:

1. To **disinfect** means to eliminate most harmful microorganisms (not including their spores) from surfaces or objects, and to inactivate viruses.
2. To **sterilize** means to kill all organism, whether harmful or not, and their spores present. In other words, sterilization is distinct from disinfection, sanitization, and pasteurization in that sterilization kills, deactivates, or eliminates all forms of life and other biological agents.

3.2.1 Physical water processes

Clarification: *A physical process is defined as a passive process that requires no energy. Alternatively, a physiologic process is defined as a living process, and it requires energy. For example, when writing with a pen, penetration of the ink into the surface/substance is a physical process, and the process of oxidation is a chemical reaction of the ink with the oxygen. Consider the settling and filtration of water for purposes of solids separation.*

Physical water processes include:

1. **Distillation -** Distillation is a process of separating the component substances from a liquid mixture by selective **evaporation** and **condensation**. It is a change of state, or a way of separating mixtures. It is the process in which water is evaporated (by an evaporator) and the vapor condensed (by a condenser). It is important to understand that evaporation is not the same as condensation although evaporation in an enclosed environment can subsequently lead to the condition of condensation as evaporated moisture is “condensed” out of the air and is reverted to a liquid stage. Distillation may result in essentially complete separation (nearly pure components), or it may be a partial separation that increases the concentration of selected components of the mixture. In either case the process exploits differences in the volatility of a mixture’s components. Distillation is a physical separation process and not a chemical reaction. The process of distillation essentially purifies water by heating it until only the water molecules evaporate and condense in another part of the filter, leaving impurities behind in the original water source. Water and dissolved salts can be separated by distillation. Scientists consider distilled water an unnaturally pure form of drinking water and as a result, substances are routinely added to remineralize it. Minerals are considered a nutritionally necessary (i.e., “healthy”) part of drinking water. Distilled water has an exceptionally low TDS, and it is often remineralized before being used as drinking water. Vapor compression distillation (technology: Slingshot water purifier). Solar distillation
2. **Alcoholic Distillation** produces a usable/drinkable concentrate of ethyl alcohol through distillation of a mixture produced from alcoholic fermentation. Distillation comes after the brewing/fermenting of a material (mixed with water) into an alcoholic solution. The product of alcoholic distillation is a water alcohol solution with disinfection, and other, properties. The distillation process purifies the solution, and removes diluting components like water, for the purpose of increasing its proportion of alcohol content (commonly expressed as alcohol by volume, ABV).
    A. **Ethyl alcohol (ethanol)**, the alcohol in alcoholic beverages, is a bacterial metabolite produced by the fermentation of sugar molecules found in various natural substances. It is a by-product of bacterial metabolism and is toxic to bacteria, as well as other organismal programs. Hence, alcohol disinfects, but also makes the result of distillation a poor hydration source. The higher the percentage of alcohol, the greater disinfection effect. Ethanol is also known as pure alcohol, grain alcohol and drinking alcohol. Ethyl alcohol is used in some cosmetics to enhance the absorption of other ingredients, which means that anything mixed with it is likely to move through the skin and into the bloodstream more easily. Ethanol is an “alcohol solvent”. Alcohol is highly flammable. Ethyl alcohol has significant and noticeable consciousness altering properties when consumed.
    B. Any alcohol when consumed in enough of a quantity can be toxic (i.e., a toxicant); however, ethanol is a “natural” by-product of bacterial fermentation and is something our body is evolved to understand. Hence, ethanol is not is not in the same toxicant category as “unnatural toxic alcohols”. “**Toxic alcohols**” are molecules which are not understood and responded to well by our human metabolism -- they cause our human physiology to systemically malfunction in various ways. They have similar harmful effects on most other organisms. This difference in relationship between categories relates to the way humans and other organisms respond when exposed to the substance [in the context of an emergently evolving biology]. To be more clear, medical practitioners refer to these alcohols as “toxic alcohols” due to their effect on human biophysiology. “Toxic alcohols” include,

but are not limited to: methanol; ethylene glycol; propylene glycol; and isopropanol (isopropyl alcohol) -- these are all "alcohol solvents" like ethanol, and hence, skin contact will lead to some absorption. In the market-State, "toxic alcohols" are more readily available than ethanol, and are usually cheaper.

C. **Denatured alcohol (SD, methylated spirits)** is ethanol that has additives to make it poisonous, extremely bad tasting, foul smelling or nauseating, to discourage recreational consumption. In some cases it is also dyed. The ethanol is made undrinkable. This alcohol is bad for your skin since it is drying and irritating, even though it has antibacterial, antiseptic, and astringent properties. In the market, denatured alcohol is used as a solvent and as fuel for alcohol burners and camping stoves. Because of the diversity of industrial uses for denatured alcohol, hundreds of additives and denaturing methods have been used. The main additive has traditionally been 10% methanol, giving rise to the term "methylated spirits". Other typical additives include isopropyl alcohol, acetone, methyl ethyl ketone, methyl isobutyl ketone, and denatonium. Methanol, also called methyl alcohol, is found in windshield wiper fluid, de-icing products, paint removers, photocopying fluid, shellacs, and embalming fluid. In humans, methanol is metabolized first to formaldehyde and then to formic acid or formate salts, which is poisonous to the central nervous system, and in sufficient concentration may cause blindness, coma, and death.

D. **Isopropyl alcohol (a.k.a., rubbing alcohol, isopropanol, Iso, IPA)** is used similarly to ethanol, except it is more toxic than ethanol, but less drying to skin. It is a petrochemical product, and numerous tests have shown it to be a strong neurotoxin (significantly stronger than ethyl alcohol). Isopropyl, like ethanol, is absorbed through the skin -- skin contact with isopropyl alcohol should be avoided. Isopropyl alcohol is a popular alternative to ethyl alcohol in commerce because it cannot be safely consumed by humans. Isopropyl alcohol cannot be used to purify water for human consumption. Isopropyl alcohol is a strong solvent and dissolves a wide range of non-polar compounds. It has the similar disinfectant properties to ethanol. It evaporates into the air quickly, and is relatively non-toxic in comparison to other petrochemical and synthesized solvents. After evaporation, isopropyl alcohol exists in the atmosphere in the gas phase. The dominant atmospheric loss process for isopropyl alcohol in the troposphere is by reaction with the hydroxyl radical. Based on this reaction, the atmospheric half-life and lifetime of isopropyl alcohol is estimated to be 1.9 days and 2.7 days, respectively (Atkinson, 1994). The major reaction product from this reaction is acetone (Atkinson, 1995). Disinfecting solutions generally consist of 60–70% solution of isopropyl alcohol in water. Water is required to open up membrane pores of bacteria, which acts as a gateway inside for isopropyl.

3. **Atmospheric water generation (AWG)** is a process that extracts water from surrounding, ambient air (i.e., from the water vapor in the air), and often, filters it to remove particulates and bacteria. The resulting water is clean and free of chemicals and other hazards. There is vapor/humidity in the air due to the evaporation of water from land. All air on Earth contains at least a little water. Hence, water generators, also known as "water makers", harvest the moisture suspended in humid air. They use one of three basic methods: water vapor in the air is condensed by cooling the air below its dew point (i.e., **condensation**); exposing the air to desiccants (i.e., **desiccation**); or pressurizing the air (i.e., **pressurization**). The most common method uses technology similar to the refrigerant-based air conditioner to condense the air into a liquid flow, filter it, and then store it in a carafe or other holding tank. The technical approach uses a water extraction method that relies on chemistry. It uses a cocktail of chemical salts to pull water out of the air. Salt is a natural desiccant, which means that it draws and holds moisture. In desiccant based water generators, humid air passes over a salt mixture. The wet salt is then heated to the boiling point. The steam is condensed and routed to filters for processing. To save on energy, a vacuum is employed to lower the boiling point of water. One big advantage to desiccant-based atmospheric water generation is that it's more energy efficient than the other practical options currently available. Typically, the reclaimed water is filtered a number of times to take out airborne particles and bacteria as part of the process. If the water sits in the reservoir for longer than a day or so, it's filtered again to keep it purified. Unlike a dehumidifier, an AWG is designed to render the water potable. If the atmosphere is contaminated, then the water produced purely from the AWG technology will be partially contaminated also. Atmospheric water generators don't work everywhere. To work

effectively, a few conditions must be met: The temperature of the ambient air has to be at least a few degrees above freezing. The humidity should be above a certain concentration, too. The figure varies depending on the manufacturer and method of extraction, but 32 to 40 percent humidity is in the ball park. High altitude can also interfere with the process. Also, because an AWG condenses water from the air it acts as a dehumidifier, and it can be combined with air-conditioning technology.

4. **Passive water generation (air well or dew harvester)** have been used for millennia. An air well or aerial well is a structure or device that collects water by promoting the **condensation** of moisture from air (Read: it uses condensation to collect moisture). Designs for air wells are many and varied, but the simplest designs are completely passive, require no external energy source and have few, if any, moving parts. Three principal designs are used for air wells, designated as high mass, radiative, and active. It appears that air wells designed to promote condensation were used by the ancient Greeks to provide water to large populations.
5. **Sedimentation** is the tendency for particles in suspension to settle out of the fluid in which they are entrained and come to rest against a barrier. This is due to their motion through the fluid in response to the forces acting on them: these forces can be due to gravity (a physical process), centrifugal acceleration (a physical process), or electromagnetism (an electromagnetic process). Controlled sedimentation (in a sedimentation basin) is when water is left undisturbed to allow heavy particles to sink out, and greases and oils to rise to the surface. The material on the top and bottom of the tanks is skimmed away. In geology, sedimentation is often used as the opposite of erosion (i.e., the terminal end of sediment transport).
6. **Filtration** is commonly the mechanical or physical operation which is used for the separation of solids from fluids (liquids or gases) by interposing a medium through which only the fluid can pass. The fluid that passes through is called the filtrate. Oversize solids in the fluid are retained, but the separation is not complete; solids will be contaminated with some fluid and filtrate will contain fine particles (depending on the pore size and filter thickness). Filter media can be cleaned by rinsing with solvents or detergents, or may be cleaned by backwashing. Self-cleaning screen filters utilize point-of-suction backwashing to clean the screen without interrupting system flow. Passing water through a bed of sand and gravel is a common last step in water purification. As water filters through the sand, the remaining particles of suspended matter are trapped in the sand bed. In the filtration process, water flows on top of the sand bed and travels through the bed until it is collected at the bottom in underdrains. Filtered water flows from the underdrains into clear wells or filtered water reservoirs.
7. **Earth filtration** are made of various mineral media, including various types of clays, diatomaceous earth, glass (silica) and other fine particles. The media are blended, shaped by manual or mechanical methods, dried and then fired at various temperatures to achieve different pore sizes and filtration properties. Some are unfired to maintain an open pore structure for filtration. Most ceramic filters are easy to use and are a potentially sustainable technology. The availability of suitable raw materials and the appropriate technology to blend these raw materials, shape the filter units and then perhaps fire them in a kiln are the main technical and accessibility barriers. The need for inspection and other quality control measures, as well as appropriate testing for proper pore size are also important requirements for their production. Some units are brittle and fragile and therefore, can break during use. Broken filters, even if only slightly cracked, are unsuitable for removal of particles and microbial contaminants from water. Well constructed earth filters have been extensively tested for efficacy in reducing various waterborne microbial contaminants. Some of these are rated to remove at least 99.9999% of bacteria, such as Klebsiella terrigena, 99.99% of viruses, such as polioviruses and rotaviruses, and 99.9% of Giardia cysts and Cryptosporidium oocysts.
8. **Ground rock filtration** is the natural filtering of particles in water by porous rocks as it moves through the ground and reappears in the form of **spring water** or **well water**. Spring water is collected immediately at the source before other particles can dissolve into it.
9. **Carbon filtration** is a method of filtering that uses a bed of activated carbon to remove contaminants and impurities, using chemical absorption. Each particle/granule of carbon provides a large surface area/pore structure, allowing contaminants the maximum possible exposure to the active sites within the filter media. One pound (450 g) of activated carbon contains a surface area of approximately 100 acres (40 Hectares). **Activated carbon** is carbon that has been processed to increase its adsorption properties. Absorption is a

process whereby [pollutant] molecules in the fluid become trapped inside the pore structure of the carbon substrate. Carbon filtering is commonly used for water and air filtration. Active charcoal carbon filters are most effective at removing chlorine, sediment, volatile organic compounds (VOCs), taste and odor from water. They are not effective at removing minerals, salts, and dissolved inorganic compounds. Typical particle sizes that can be removed by carbon filters range from 0.5 to 50 micrometres. The particle size will be used as part of the filter description. The efficacy of a carbon filter is also based upon the flow rate regulation. When the water is allowed to flow through the filter at a slower rate, the contaminants are exposed to the filter media for a longer amount of time.

10. **Membrane filtration** (ultra/micro/nanofiltration) is a water separation process. Water processing membranes are thin sheets of material that are able to separate water from other elements based on properties such as size or charge. Water passes through a membrane; but depending on their size, larger particles, microorganisms, and other elements are separated out. Some of these systems are pressure driven and depend on water pressure to separate the particles based on size. Reverse osmosis and nanofiltration are two membrane technologies. Membranes are often made of plastic or ceramic, with tiny openings called pores through which the water is filtered. The smaller the pore, the greater quantity of pollutants rejected and the more energy needed to force the water through the membrane. Membrane filtration produces clean water ("permeate") and a "waste" stream (concentrate/retentate) is discharged that contains all of the material that has been filtered out. The higher the TDS of water, the more difficult it becomes to maintain a membrane filter. Membrane filter types include:
    A. **Hyper filtration (HF)** can separate larger ions and molecules from smaller ones. (0.0001 micron pore size)
    B. **Ultra filtration (UF)** using hollow fibres can filter the smallest bacteria. (0.01 to 0.1 micron)
    C. **Nano filtration (NF)** use membranes to filter viral organisms. (0.001 to 0.01 micron)
11. **Reverse osmosis (RO) membrane filtration** is a physical membrane process for separation of dissolved substances into two streams, permeate (purified) and concentrate (unpurified). Essentially, reverse osmosis (RO) is a physical process for the de-concentration of substances dissolved in liquids. The process occurs by drawing water through a fine membrane under pressure. While the process is comparatively slow, this type of purification system can rid water of almost any contaminant, including radioactive particulates. The filtered water is stored in a closed tank to prevent contamination. The RO process strips the water down to mostly H20 molecules, and hence, minerals are routinely re-added to improve the quality of the water (if it is to be used for drinking water). The pore size of reverse osmosis membranes is so small that mostly water passes through. The filtration generally begins with the water flowing first through an activated charcoal pre-filter where it is cleansed of large molecular materials. Next the water flows through the main RO filter (i.e., screen/diaphragm). The tiny water molecules easily pass through the half permeable ("semi-permeable") diaphragm, which functions like a screen with an extremely small mesh size (2-3 angstroms). Nearly, but not all, pollutants are substantially larger in their molecule size than the H2O (water) molecule; thus, the pollutants get caught in the reverse osmosis diaphragm. These filtered pollutants are then rinsed off by the next flow of water, and washed down the concentrate drain. In this way the diaphragm is constantly cleaned and prevented from accumulating pollutants. Reverse osmosis systems produce larger quantities of wastewater (or "concentrate") than most other systems. They can become clogged with clay or organic materials if particle-rich source water is not filtered first. Maintenance is generally not difficult, but it can be expensive since the main action required is to replace the membrane as necessary. Maintenance problems tend to involve leaking and fouling of membranes.
12. **Membrane bioreactors (MBR)** use a water purification process that integrates a perm-selective or semi-permeable filtration membrane with a biological process contained within a bioreactor. The anaerobic membrane bioreactor (AnMBR), which is a combination of the anaerobic biological wastewater treatment process and membrane filtration, represents a recent development in the high-rate anaerobic bioreactors.

3.2.2 Biological water processes

Biological water processes include:

1. **Bacterial processes** are introduced, and they begin consuming the small particles of organic matter that remain in the water. To facilitate this, oxygen is pumped into the water to allow the bacteria to feed continuously, as low oxygen levels cause feeding

rates to slow.

2. **Myco-filtration** as **mycological processes** is a biological approach to water purification that uses the web-like tissue of mushroom-forming fungi (i.e., mycelium) to capture and degrade environmental pollutants. Mycofiltration is a sustainable approach to storm water processing.
3. **Slow sand filters** are large open sand beds without any backwashing mechanism. They are excellent at removing coliform bacteria and protozoans such as Giardia and Cryptosporidium. A slow sand filter not only physically filters the water, but also provides biological treatment from an organic mat which naturally forms on the filter surface. A slow sand filter can reduce microbial contaminants without the high cost of coagulation or pre-processing chemicals needed for common high-rate filters. Slow sand filtration percolates source water slowly through a bed of porous sand, with the influent water introduced over the surface of the filter, and then drained from the bottom. Properly constructed, the filter consists of a tank, a bed of fine sand, a layer of gravel to support the sand, a system of underdrains to collect the filtered water, and a flow regulator to control the filtration rate. No chemicals are added to aid the filtration process. The slow sand filter reduces suspended organic and inorganic matter, and may remove some pathogenic organisms. A granular activated carbon (GAC) sandwich filter is a modified slow sand filter that removes organic material. The removal action includes a biological process in addition to physical and chemical ones. Slow sand filters are less effective at removing microorganisms from cold water because as temperatures decrease, the biological activity within the filter bed declines. Since the purification mechanism in a slow sand filter is essentially a biological process, its efficiency depends upon a balanced biological community in the schmutzdecke. A sticky mat of biological matter, called a "schmutzdecke," forms on the sand surface, where particles are trapped and organic matter is biologically degraded. Slow sand filters rely on this cake filtration at the surface of the filter for particulate straining. As the surface cake develops during the filtration cycle, the cake assumes the dominant role in filtration rather than the granular media. A slow sand filter must be cleaned when the fine sand becomes clogged, which is measured by the head loss. The length of time between cleanings can range from several weeks to a year, depending on the raw water quality. The operator cleans the filter by scraping off the top layer of the filter bed. A ripening period of one to two days is required for scraped sand to produce a functioning biological filter. The filtered water quality is poor during this time and should not be used. Slow sand filter monitoring and operation is not complicated. Daily tasks include reading and recording head loss, raw and filtered water turbidity, flow rates, and disinfectant residual. If necessary, the operator should adjust the flow to bring water production in line with demand.

3.2.3 Chemical water processes

A chemical reaction is a process that leads to the transformation of one set of chemical substances to another.

Chemical water processes include:

1. **Coagulation** and **Flocculation** - Particles suspended in water naturally repel other particles. They cannot join to form larger particles that would settle more quickly. Coagulation occurs when a coagulant is added to water to "destabilize" colloidal suspensions. A coagulant, such as alum, is added to the source water to overcome the repulsion between the particles. Flocculation involves gently mixing the water so that the particles can combine and form larger particles. Flocculation, in the field of chemistry, is a process wherein colloids come out of suspension in the form of floc or flake; either spontaneously or due to the addition of a clarifying agent. The action differs from precipitation in that, prior to flocculation, colloids are merely suspended in a liquid and not actually dissolved in a solution. In the flocculated system, there is no formation of a cake, since all the flocs are in the suspension. These processes are essential pre-processing for many water purification systems. In conventional coagulation-flocculation-sedimentation, a coagulant is added to the source water to create an attraction among the suspended particles. The mixture is slowly stirred to induce particles to clump together into "flocs." The water is then moved into a quiet sedimentation basin to settle out the solids. Dissolved air flotation systems also add a coagulant and flocculate the suspended particles; but instead of using sedimentation, pressurized air bubbles force them to the water surface where they can be skimmed off. A flocculation-chlorination system has been developed as a point-of-use technology. It uses small packets of chemicals and simple equipment like buckets and a cloth filter to purify the water.

Finally, lime softening is a technology typically used to “soften” water—that is, to remove calcium and magnesium mineral salts. In this case, the material that is settled out is not suspended sediment but dissolved salts.

2. **Ozone (O_3)**, or trioxygen, is an inorganic molecule with the chemical formula O_3. It is an unstable form of pure oxygen (0_2); it is a gas. Ozone the tri-atomic form of oxygen: instead of the normal arrangement of 2 atoms of oxygen (O_2), ozone is comprised of 3 atoms of oxygen (O_3). Ozone is nothing other than ionized oxygen. Ozone, however, doesn’t want to stay in that tri-atomic state very long and unless held in check or bound by other molecular couplings, ozone will usually break down from O_3 to O_2 + O_1 within 20 minutes (at approximately atmospheric pressure). O_1 is called a singlet oxygen atom and it is highly reactive - it is a powerful oxidant with bactericidal properties very similar to chlorine. Because ozone is highly reactive, it readily oxidizes (i.e., breaks down) organic matter. When ozone encounters another compound, one oxygen atom will break away, attach itself to the compound, and oxidize (“clean” or “purify”) it. Ozone quickly reverts back into oxygen, leaving nothing behind other than oxygen. In concentration, it will reduce or prevent the growth of micro-organisms including bacteria, viruses, cysts, spores, mold and mildew. Ozone is also effective in eliminating or controlling color, taste, and odor problems. It oxidizes iron and manganese into solid particles that can be filtered out. Ozone is produced in nature when the ultraviolet rays of the sun strike oxygen molecules. When this occurs stratospheric layer of the atmosphere, it creases what we know as the “ozone layer”. Nature also produces ozone near the earth’s surface. Ozone is created when lightening, air, and precipitation combine. The sweet, fresh smell after a thunderstorm is the smell of ozone gas. Its efficacy is not dependent upon pH as is that of chlorine. There are 3 common methods of producing ozone: hot spark as corona discharge (CD), ultraviolet light, and cold plasma. A typical ozone system consists of an ozone generator mounted on the wall of a storage tank. The system uses a pump to circulate the ozone through the water. And, a filter is generally placed inside an storage tank to catch particles. Rising bubbles of ozone and air cause water to be drawn through the filter and it cleans the water without a water pump. The ozonation process is currently the most effective, sustainable, and safest non-chemical method of eliminating bacteria from water. Due to the highly reactive nature of ozone, it can corrode some materials. Hence, all surfaces coming in contact with ozone should be made of ozone-resistant materials, such as stainless steel. Ozone is an irritant (i.e., caustic) to lung and nasal tissue at concentration, and hence, generators may leak and could create an ozone hazard. However, ozone dissolved in water will not irritate skin, nose, or ears, nor will it dry out or leave a chemical film on skin. Ozone does not affect the pH balance of water like traditional chemical treatment methods. Ozone is a powerful sanitizer and oxidizer and reverts back into oxygen, leaving no harmful by-products whatsoever. Ozone can be used to purify a water body, including water in containers, storage tanks, spas, bathtubs, fountains, and wading pools. An ozonation system will sanitize the tank as well as the air above the water line. Most ozone systems require a storage tank for the treated water. With ozone treatment, disinfection occurs primarily at the point of contact between the ozone and the water. The disinfection process does not occur beyond the treatment unit. This contrasts with chlorination treatment where the residual chlorine remains in the water and continues the disinfection process for some time. Because ozone is so unstable, it does not produce a reliable residual. Ozone has an active residual time measured in minutes, whereas the active residual time for chlorine is measured in hours. Here, temperature is an important variable. The colder the water the more ozone will be dissolved in water and the longer the residual. Room temperature water looses its concentration of ozone within several minutes. Ozone is continuously moving out of solution and will dissipate completely from very cold water (context dependent) within about 20 minutes. In general, the lower the water temperature, the more ozone is dissolved in the water. Ozone has been found to be effective over a wide range of pH, but a pH slightly above 7 increases treatment efficiency. The ozone demand is related to the level of contamination in the water. When substances in the untreated water react with ozone, part of the ozone is used up, which may leave less ozone available to treat the targeted contaminants. Ozonation units are installed as a point-of-entry treatment system. At the point where the ozone mixes with the water, turbulence and bubbles are created (through aerated diffusion of the ozone); these ensure that the ozone contacts as much of the raw water as possible, and is dissolved into the water to some degree. The greater the water flow rate, aeration (e.g., “diffuser stone”), and pressure differential/turbulence, the

more effective the “treatment”. The effectiveness of ozonation depends significantly on the contact time, temperature, and solubility. In general, ozone requires a shorter contact time than chlorine. Ozone treatment can produce harmful by-products in water. For example, if bromide is present in the raw water, ozone reacts with it to form bromate, shown to cause cancer in rats. The less humidity an ozone generator has to work with, the better the results. The perfect environment for ozone production is close to 15-38°c with no humidity. Ozone can still be made in an environment with high humidity (50-80%), but with an air drier, the ozone generator will double or even triple the amounts of ozone produced, cleaning your water in less time. An air drier will also lengthen the life of the ozone machine by allowing it to run more efficiently. After a period of time, an ozone machine (without an air dryer) converts moisture to an acidic yellow coloured paste. This paste build-up will ruin the ozone generator causing poor ozone performance. Most ozone systems do not require extensive maintenance. Some systems use an air-drying material, which needs to be replaced periodically. It is also necessary to periodically clean the water storage tank and check pumps, fans and valves for damage and wear. If UV radiation generates the ozone, the lamp must be replaced periodically. Any pre-treatment or post-treatment devices may require additional maintenance. Additional uses for ozone: Ozonated water can be used to sanitize surfaces when submerged in a sufficient concentration for a sufficient amount of time. In other words, it can be used to disinfect toothbrushes and dentures. It can be used as a gargling solution. I can be used as type of “air purifier”.

3.2.4 Electromagnetic processes

Electromagnetic water processes include:

1. **Ultraviolet (UV) radiation/energy** is found in the electromagnetic spectrum between visible light and x-rays and can best be described as invisible radiation with disinfection as a property. The ultraviolet bandwidth occupies wavelengths roughly between 200 and 400 nanometers. Purification via exposure to ultraviolet radiation (UV germicidal irradiation) is unique from other types of sterilization modalities due to the fact that it does not necessarily cause death of the target organism. In those pathogens it does not directly kill, the UV radiation effectively alters the organism’s genetic structure. UV energy penetrates the outer cell membrane, passes through the cell body and disrupts its DNA preventing reproduction. By causing damage to the target bacteria’s Deoxyribonucleic Acid (DNA), the bacteria is sterilized at the genetic level. Thus, the organism is no longer able to reproduce and cause disease. The degree of inactivation by ultraviolet radiation is directly related to the UV dose applied to the water. The dosage, a product of UV light intensity and exposure time, is measured in microwatt second per square centimeter ($\mu ws/cm^2$). UV treatment does not alter water chemically; nothing is being added except energy. UV units only kill bacteria at one point in a watering system and do not provide any residual germicidal effect downstream. The sterilized microorganisms are not removed from the water -- bacteria cells are converted into pyrogens. The killed microorganisms and any other contaminants in the water are a food source for any bacteria that do survive downstream of the UV unit. Due to these limitations, the piping in a watering system treated by UV disinfection will need to be periodically sanitized with a chemical disinfectant. UV disinfection does not remove dissolved organics, inorganics or particles in the water. Although 100% destruction of microorganisms cannot be guaranteed, it is possible to achieve 99.9% reduction in certain applications and with proper maintenance. In order for a UV unit to successfully disinfect water, the following additional variables must be considered: Firstly, the composition of water can reduce the transmission of UV light through the water, which reduces the UV dose that reaches the bacteria. UV disinfection is most effective for treating high-clarity purified reverse osmosis or distilled water. Secondly, all UV units have a maximum flowrate capacity and some have a minimum flowrate as well. Finally, UV radiation generates ozone that can damage surrounding materials.
2. **Boiling** is the oldest water disinfection technique. Boiling is the rapid vaporization of a liquid, which occurs when a liquid is heated to its boiling point, the temperature at which the vapor pressure of the liquid is equal to the pressure exerted on the liquid by the surrounding environmental pressure. effective despite contaminants or particles present in it, and is a single step process which eliminates most microbes responsible for causing intestine related diseases. In places having a proper water purification system, it is recommended only as an emergency treatment method or for obtaining potable water in the wilderness or in rural areas, as it cannot remove chemical toxins or impurities.

3. **Steam and pressure sterilization** through the use of an autoclave. An autoclave is a pressure chamber used to sterilize materials by subjecting them to high pressure saturated steam for a duration of time.

3.2.5 Electrochemical water processes

Electrochemical approaches to water tend to emphasize the recovery of metals and chemicals, and the destruction of organic contaminants.

Electrochemical water processes include:

1. **Desalination** is the process of removing dissolved salts and minerals from seawater or brackish water. It is also called desalination or desalting. Desalination is an advanced technical process involving many different technologies, including distillation, reverse osmosis, electrochemical mediation, and membrane filtration. Desalination produces drinking water and concentrate (the water that contains the salts that were removed in the desalination process, which used to be called brine). Salt water is desalinated to produce fresh water suitable for human and ecological usage. The minerals in the concentrate can be separated to produce, at the very least, a large amount of salt and a complex mineral solution for the biological nutrition needs of various organisms, notably, vegetation In other words, ocean water can be desalinated to produce a complex mineral solution (i.e., nutrient source) for plants and other organisms.
2. **Electrodialysis** is a water processing system that uses electrical potential to remove charged impurities from water. It is widely used for the processing of brackish water and for the desalination of seawater. Unlike other membrane processes, the source water never passes through the membranes during electrodialysis. The migration of ions across the membrane under the influence of a potential gradient is the basis of electrodialysis. Early processes used a three-compartment cell and the membranes were non-selective. Today, an array of membranes are used with alternating cation-selectivity and anion-selectivity.
3. **Disinfection** is the chemical or biological destruction of pathogenic microorganisms such as viruses, protozoa, and bacteria, or at least the deactivation of those microorganisms. Several chemicals are used in water disinfection and electro-chemistry is involved in the production of those chemicals. Common disinfection chemicals include: halogens; chlorine (Cl); chloramine (NH_2Cl); chlorine dioxide (ClO2); bromine (Br); hypochlorite (ClO-); ozone (O^3); hydrogen peroxide (H202); and iodine (I). The importance of these chemicals is based predominantly on their oxidizing power. They provide a flexible, controllable bioaction that can be maintained throughout a water distribution system. However, these chemicals have varying residual disinfection properties.

A. **Ozone (O3)** is a strong oxidizing substance with bactericidal properties very similar to chlorine. It is claimed that ozone destroys bacteria more rapidly than chlorine and its efficacy is not dependent upon pH as is that of chlorine. The ozonation process is currently the most effective, sustainable, and safest non-chemical method of eliminating bacteria from water. Note, the major difference between ozone and UV is that ozone is a very powerful oxidizer and kills all known microorganisms—whereas UV typically just inactivates them. Ozone actually opens up the cell wall of the bacteria, virus, mold, fungus, etc. Then, because ozone oxidizes them (opens up the cell wall), the impurities clump together and are more easily filtered out. Water becomes clear with ozone, and the clearer the water is, the more effective UV is. Only clear water can be effectively treated with UV, cloudiness in the water can absorb the UV light. If you use UV, you should definitely use ozone too. But if you use ozone—you don't necessarily need UV. Ozone does everything UV does and more. But UV does not do everything that ozone does.

B. **Hydrogen peroxide** is a chemical compound with the formula H_2O_2**.** Hydrogen peroxide should really be called hydrogen dioxide. In its pure form, it is a colourless liquid, slightly more viscous than water; however, for safety reasons it is normally used as an aqueous solution. When exposed to other compounds hydrogen peroxide dismutates readily. The extra oxygen atom is released leaving H_20 (water). In nature oxygen (0_2) consists of two atoms--a very stable combination. H_20_2 is found naturally in rainwater. There is ozone (O_3) in the atmosphere, and when it rains, the unstable ozone gives up atoms of oxygen to falling water to form hydrogen peroxide. It turns out that this is one of the reasons that rainwater causes plants to grow more rapidly. In the human body, the cells responsible for fighting infection and foreign invaders in the body (your white blood cells) make hydrogen peroxide and use it to oxidize pathogens. Hydrogen peroxide is the simplest peroxide (a compound with an oxygen–oxygen single bond) and finds use as a strong oxidizer, bleaching agent and disinfectant. When stored under the proper conditions, it is a very stable compound. Concentrated hydrogen peroxide, or "high-test peroxide", is a reactive oxygen species and has been used as a propellant in rocketry. Pure hydrogen peroxide will explode if heated to boiling. Hence, for general usage, it is usually handled as a dilute solution. 3–6% H_2O_2 will not burn skin, but will kill bacteria on a wound. 32-35% (a.k.a., food grade) will burn skin on contact and is generally diluted before use. On human skin, peroxide reacts with melanin to oxidize it and convert it into a compound which is free of colour. Hydrogen peroxide's same oxidizing properties allow it to react with bacteria, viruses, spores, and yeasts, making it a functional disinfectant. It is rarely used in drinking water production as a stand-alone treatment process. Hydrogen peroxide solution may be used to disinfect surfaces. When it comes into contact with bacteria on a wound there is a fizzing and bubbling that happens, which is the result of the H_20_2 bonds breaking during the reaction. One of the oxygen molecules is liberated, leaving H_20. The free oxygen molecules are what you see bubbling to the surface. It has a variety of medical uses including that of mouthwash (wherein it whitens teeth) and used to treat fungal issues on tissue. Hydrogen peroxide is odorless and colorless, but not tasteless. It can also be used in the ear to treat ear infections. It can be used to disinfect surfaces like toothbrushes and countertops. It can be used as a bleach on skin, grout, hair, and clothes. It can be added to a clothes washing agent to facilitate cleaning and to bleach fabric. It can be used on vegetables and will disinfect/ sanitize utensils. Plant seeds can be soaked in a low-grade hydrogen peroxide solution to simulate rainwater. It can be used in aquatic environments to reduce algae.

C. **Chlorine** (and **chloramine**) is used to kill any harmful bacteria that might be present in a water source, and it is the most widely use disinfecting agent in early 21st century society. The amount of chlorine that is used is based on the amount of water that is treated, the amount of contaminants that must be controlled, and the time it takes for the water to reach its first usage (i.e., residual disinfection).

4. **Bleaching** refers to the usage of a number of chemicals which remove color, whiten or disinfect, often by oxidation. Chlorine is the basis for the most commonly used bleaches. Bleach performs two types of actions: it whitens, and has antimicrobial efficacy. There are two primary classes of bleach: chlorine-based bleach and peroxide-based bleach. Sodium hypochlorite is the most commonly encountered bleaching agent, usually as a dilute (3–6%) solution in water. Chlorine bleach forms dioxin after contact with organic compounds. Dioxins are a group of chemically-related compounds that are persistent environmental pollutants (POPs). Dioxins are highly toxic and can cause reproductive and developmental problems, damage the immune system, interfere with hormones and also cause cancer.

3.2.6 Ecological processes

Consider the role of water in the structure and function of ecological processes.

1. **Waste Stabilization Ponds (WSP) and Constructed Wetlands (CW) (**or **CWWSP)** are a form of wastewater and stormwater processing, which involve the construction of low energy-consuming ecosystems that use natural processes to transform, decompose, and recycle materials. A CWWSP forms a controlled ecological network of ponds (Read: a pond system) for the transformation, decomposition, and recycling of materials from used water and storm water. Together, the different ecological compositions of the ponds in the network, and the flow of water therein, form a unified ecological transformation system for processing water with a portion of the material effects of our living into sustainable building blocks (i.e., nutrients) for further life. Therein, interconnected pond water levels dictated by ground water levels. The pond system attract wildlife. It is essentially a system designed to mimic (as in, biomimicry) the way nature transforms water.
2. **Waste Stabilization Ponds (WSP)** are large, shallow basins in which raw sewage is processed (i.e., "treated") entirely by natural processes, primarily involving algae and bacteria.
   A. **Anaerobic ponds** - Anaerobic ponds are commonly 2 – 5m deep and receive wastewater with high organic loads (i.e., usually greater than 100 g biochemical oxygen demand (BOD)/m3 .day, equivalent to more than 3000 kg/ha.day for a depth of 3m). They normally do not contain

dissolved oxygen or algae. In anaerobic ponds, BOD removal is achieved by sedimentation of solids, and subsequent anaerobic digestion in the resulting sludge.

B. **Faculative ponds** - (1-2m deep) are of two types: Primary facultative ponds that receive raw wastewater, and secondary facultative ponds that receive particle-free wastewater (usually from anaerobic ponds, septic tanks, primary facultative ponds, and shallow sewerage systems). The process of oxidation of organic matter by aerobic bacteria is usually dominant in primary facultative ponds or secondary facultative ponds.

C. **Maturation ponds** are usually 1-1.5m deep, receive the effluent from the facultative ponds. Their primary function is to remove excreted pathogens. Although maturation ponds achieve only a small degree of BOD removal, their contribution to nutrient removal also can be significant. Maturation ponds usually show less vertical biological and physicochemical stratification, and are well-oxygenated throughout the day. The algal population in maturation ponds is much more diverse than that of the facultative ponds, with non-motile genera tending to

3. **Constructed wetlands (CW)** are planned systems designed and constructed to employ wetland vegetation to assist in processing wastewater in a more controlled environment than occurs in natural wetlands, but mirroring the way the natural wetlands function. Herein, aquatic plants as a bio-filter/bio-transformer. Essentially, a constructed wetland is a designed complex of saturated substrate, emergent and submerged vegetation, animal life, and water that simulate wetlands for human uses and fulfillment. In early 21st century society, constructed wetlands are an "eco-friendly" alternative for secondary and tertiary municipal and industrial wastewater treatment. The pollutants removed by CW's include organic materials, suspended solids, nutrients, pathogens, heavy metals and other toxic or hazardous pollutants. In municipal applications, they can follow traditional sewage treatment processes. Different types of constructed wetlands can effectively treat primary, secondary or tertiary treated sewage. However wetlands should not be used to treat raw sewage and, in industrial situations, the wastes may need to be pre-treated so that the biological elements of the wetlands can function effectively with the effluent. CW's are practical alternatives to conventional treatment of domestic sewage, industrial and agricultural wastes, storm water runoff, and acid mining drainage.
   - Constructed wetlands for wastewater treatment can be categorized as either **Free Water Surface (FWS)** or **Subsurface Flow (SSF)** systems. In FWS systems, the flow of water is above the ground, and plants are rooted in the sediment layer at the base of water column. In SSF systems, water flows though a porous media such as gravels or aggregates, in which the plants are rooted.
4. **Riverbank filtration (bank filtration)** - The improvement of water quality through natural filtration by soil and microbial processes caused by the percolation of water through natural or engineered media. Bank filtration involves installing a well near the surface water source to allow pulling water through the ground to effect some natural filtration. This concept requires granular soils adjacent to or under the surface water source. Where these conditions exist, there can be significant benefits with employing bank filtration.
5. **Reed beds** are natural habitats found in floodplains, waterlogged depressions and estuaries. Reed beds are part of a succession from young reed colonising open water or wet ground through a gradation of increasingly dry ground. As reed beds age, they build up a considerable litter layer which eventually rises above the water level, and ultimately provides opportunities for scrub or woodland invasion. Artificial reed beds are used as a method of removing pollutants from grey water, which will concentrate many of the undesirable elements of water into the stems of the plants given sufficient time. The Common Reed (Phragmites Australis) has the ability to transfer oxygen from its leaves, down through its stem, porous speta and rhizomes, and out via its root system into the rhizosphere (root system.) As a result of this action, a very high population of micro-organisms occurs in the rhizosphere, with zones of aerobic, anoxic, and anaerobic conditions. Therefore with the wastewater moving very slowly and carefully through the mass of reed roots, this liquid can be successfully treated, in a manner somewhat similar to the conventional biological filter bed systems of sewage treatment plants.

3.2.6.1 Additional pond functions in a pond system

There are many practical uses for a pond. Below are some of the other ponds in a pond system.

> **NOTE:** *Ponds cannot be discharged for aesthetic reasons.*

1. **Evaporation ponds** evaporate water from a surface area by sunlight and exposure to the ambient temperatures. Evaporation ponds have several uses. They can be used to evaporate stormwater. They are also used to dispose of brine from desalination plants. Mines use ponds to separate ore from water. Evaporation ponds at contaminated sites remove the water from hazardous waste, which greatly reduces its weight and volume and allows the waste to be more easily transported, treated and stored. They can be used to produce salt and mineral solutions from sea water (i.e., salt evaporation ponds). However, any contaminants in the seawater will end up resulting solution. The evaporation of water leaves most, but not all, salts and impurities behind. In the agricultural industry, evaporation ponds are used to prevent pesticides, fertilizers and salts from agricultural wastewater from contaminating the water bodies they would flow into.
2. **Irrigation ponds** store and supply water for later usage in irrigation. Ponds can also be used to regulate stream flow. Overhead sprinkler often lead to a lot of water evaporation prior to reaching the soil.
3. **[Storm] Drainage ponds** and **retention basins** for detention/retention/evaporation of storm and flood water. Storm drainage ponds are part of the overall surface water management system on a site and are used for flood control. The purpose is to store water during peak storm events to slow the discharge offsite and contribute to the flood control of a local storm area. There are two basic types of storage: an above ground pond or underground structures of vaults or high-volume pipes. Ponds generally have an"outflow structure" composed of an inlet and a metered outflow structure. Surface ponds are usually planted with vegetation that serves as a pollutant filter of water before being discharged downstream.
4. **Swimming ponds** are used for human bathing and swimming.

3.2.6.2 Natural swimming ponds/pools

A natural swimming pool or natural swimming pond (NSP) is a system consisting of a constructed body of water, where the water is contained by an isolating membrane or membranes, in which no chemicals or devices that disinfect or sterilize water are used, and all clarifying and purifying of the water is achieved through biological filters and plants rooted hydroponically in the system.

It is called a "natural swimming pool" because the filtration systems used have biological equivalents in the natural world. In fact if there is not an example of the type of filtration being used in the natural world then it is not seen by the IOB (International Organization for natural Bathing waters) to be naturally filtered. Natural aquarium and koi ponds also exist, but use algae scrubbers for their filtration. Systems that use UV, ozone and copper/silver ion disinfection techniques are examples of additive methods.

The NSP is divided into two areas, the swimming zone and the regeneration zone, which are physically separated. The regeneration zone and swimming zone must be equal in area for sufficient purification. The swimming portion of the pool can look like a conventional swimming pool or a natural pond. The regeneration zone and swimming zone must be equal in area for sufficient purification. The swimming portion of the pool can look like a conventional swimming pool or a natural pond. The regeneration zone can be placed adjacent to the swimming area or in a remote location depending on the space available. In up-to-date natural swimming pools there is no minimum depth for the swimming zone and the regeneration zone can now be reduced greatly and in some cases is non existent.

NSPs rely on a combination of hydraulic design techniques in conjunction with a fine filter. They require a specialized skimmer or overflow channel. Downstream, a bioreactive biological filter removes and retains additional particulate matter down to a particle size of 100 micrometres. Pumps different from those found in conventional pools hydraulically optimize water flow rates and volumes, thereby accelerating the cleaning process even further. This mechanically enhanced natural filtration produces clear, clean water. In fact, when NSPs are installed in Europe, pool owners and builders will ceremoniously wrap up an installation with a ceremonial drink of water collected directly from the pool. A biofermenta system is the latest form of filtration for NSPs reducing the need for large plantation areas and filtration beds.

In an NSP, swimming pool water flows via gravity from the swimming pool into distribution shafts. The water then passes through a biological fine filter/bioreactor before it is pumped into the regeneration zone. Plants in the regeneration zone are planted in the substrate, and these plants compete for nutrients that would feed algae. The water is cleaned biologically by the metabolic processes found in the substrate in which the plants are rooted. Thus, microorganisms and the water plants ensure effective, continuous cleaning. No processes beyond these, such as UV sterilization, are necessary unless something is added into the system that causes an imbalance.

Because of their reliance on natural environmental factors, each NSP system is built with consideration to the region and climate where it is installed. Construction elements such as the biological filter and the combination of contaminant ridding plants vary with each pool. These ponds may also be connected to a larger pond water filtration, purification, and recycling system.

Of note, natural swimming ponds can be heated to

27°c, and under certain circumstances, higher than that. These ponds do require some additional pipework and a heating system. Once people become acclimated to swimming in a pond, temperature is often not found to be the issue they thought it was beforehand. Particularly for natural designs, water in the shallow areas heats up quickly in the sun and can raise the natural temperature of the whole pond to 75 degrees during the summer in some climates. Heat pumps and solar panel warming systems can easily be incorporated into NSP construction.

Perhaps more importantly, natural water “feels” much more comfortable on the skin. It has a different composition of suspended and dissolved solids, and so, what might feel cold in chlorinated water, 71 degrees for example, actually feels comfortable in natural water.

Typically, a chemically treated pool can discharge up to 3 times its volume of water into the sewer per year. As there is no use of chemicals in the water of NSPs and their water is maintained as “living” (i.e. the NSP is not drained and refilled), there is no need for this waste of water being discharged into the sewer. Also, there is considerably reduced energy consumption for the mechanical operation of many NSP systems.

Each NSP is unique, and so each has a different “break-in” period during which the system finds balance. While NSPs can be used for swimming immediately, it takes approximately two to three years until a stable biological equilibrium is reached. Once an NSP has reached equilibrium, it requires considerably less regular maintenance than a conventional swimming pool. Seasonal care and maintenance is still required over the course of the year, though even this differs from care of a conventional swimming pool. For example, ducks and other water fowl in migration have been known to visit NSPs during the winter months. Additionally, amphibious and aquatic creatures like frogs, salamanders, and snails often make their homes in the regeneration zones of NSPs.

3.2.6.3 Additional benefits of ponds

Besides water transformation ponds provide a number of other benefits.

> **NOTE:** *Ponds do not function well on sandy or other highly porous soils*

1. Ponds are often used for frost protection, particularly on wine grapes. The use of water for this purpose typically ranges from 0.4-1.6 inches of water in a year.
2. Ponds can be managed to provide wildlife habitat.
3. Ponds constructed primarily for fish production, typically at least a half-acre in size and a minimum depth of 8 feet, can yield 100-300 pounds of fish per year for each acre of water surface.
4. Ponds can assist in flood control by capturing and slowing the flow of water through a watershed.
5. Ponds help recharge groundwater. Whether filled with water diverted from a stream or with tailwater from irrigation, clay-lined ponds seep water into the ground at highly variable rates (depending on size and construction), but typical seepage loss from a well-sealed pond is estimated at one foot of water per year.
6. Storing water captured from rainfall or from tailwater in on-farm ponds can also reduce energy use by displacing pumped groundwater.
7. Ponds can serve as water sources for fire protection if they are sited in proximity to structures.
8. Ponds can be used to settle and filter runoff, capturing soil that can be returned to use and filtering pollutants and particulates that would otherwise negatively impact the broader ecosystem.
9. A pond network creates a more localized and distributed water supply, which can offset water transported from distant reservoirs, reducing the energy needed for water conveyance.
10. Ponds can be used in the intentional and directed flooding application known as rapid flood flow irrigation. Rapid flood flow irrigation involves the application of a large amount of water to the land, which soaks in, and does not have significant evaporation.

4 Water heating

Water heating is a thermodynamic process that uses an energy source to heat water above its initial temperature. Heated water is used for cooking, cleaning, bathing, space heating. In materialization and other services, hot water and water heated to steam have many uses. Water heating systems are categorized by their energy source and whether the system is tank-type or tankless.

- Tank-based water heating system - either the storage tank is directly heated, or panels that cycle water through the storage tank are heated with an energy source such as solar, geothermal, or wood.
- Instantaneous water heating systems provide hot water on-demand at their point-of-use (POU). These systems may use a water storage tank, but the water in the tank is not hot; it becomes hot at its point of use.

5 Cleaning

Regular cleaning of material surfaces reduces the build-up of environmental materials and decomposition factors, while also reducing the amount of organic (and inorganic) matter that contributes to the proliferation of bacteria and viruses, and may inhibit the intended functioning of the surface. The purpose of cleaning is to remove the build-up of materials and to prevent future accumulation. Cleaning is essential for food production and preservation, as well as for the care of textiles. Cleaning must be considered in the context of an ecology, in particular, a bacterial ecology.

1. Physical / mechanical pressure cleaning - a process that utilizes mechanical energy (e.g., washing machine, pressure washing, scrubbing).
2. Ultrasonic cleaning - a process that utilizes sound energy to create cavitation of microscopic bubbles on the surface of an object.
3. Chemical cleaning - a process that utilizes a chemical agent (i.e., soaps/detergents).

The performance of a “cleaner” can be measured via the following indicators:

1. If a product leaves any residue from the detergent solution, then it is not clean.
2. If a product adds to the weight of the material, then it is not clean.
3. If the cleaning damages the surface unnecessarily, then it is a poor quality cleaner.

5.1 Mechanical pressure washing

Pressure washing uses a high-pressure mechanical sprayer used to remove material from surfaces and objects such as buildings, vehicles, and concrete with water [under pressure]. They are used primarily as exterior cleaning tools. The volume of a pressure washer is expressed in litres per minute, and may be variable or static. The basic pressure washer consists of a motor (either electric, internal combustion, pneumatic or hydraulic) that drives a high-pressure water pump, a high-pressure hose and a trigger gun-style switch. Just as a garden hose nozzle is used to increase the velocity of water, a pressure washer creates high pressure and velocity.

Some washers, with an appropriate nozzle, allow detergent to be introduced into the water stream, assisting in the cleaning process. Different pressure washers can use water at differing temperatures. And, there are different types of nozzles for different applications.

Washers are dangerous tools and should be operated with due regard to safety instructions. The water pressure near the nozzle is powerful enough to strip flesh from bone. Particles in the water supply are ejected from the nozzle at great velocities. The cleaning process can propel objects dislodged from the surface being cleaned, also at great velocities. Pressure washers have a tendency to break up tarmac if aimed directly at it, due to high-pressure water entering cracks and voids in the surface. Washers can damage surfaces: water can be forced deep into bare wood and masonry, leading to an extended drying period. Such surfaces can appear dry after a short period, but still contain significant amounts of moisture that can hinder painting or sealing efforts.

5.2 Cleaning instruments

Cleaning instruments include, but are not limited to:

1. A **sponge** is a tool or cleaning aid consisting of porous material. Sponges are usually used for cleaning impervious surfaces. They are especially good at absorbing water and water-based solutions. Natural sponges are generally known as a "loofah". A natural loofah is a plant seed pod in the cucumber or gourd family that grows on a vine. The pod is rough textured when dry, and softens when combined with water.

5.3 Cleaning with water 'washing agents'

The two primary types of washing agents are soaps and detergents. The distinction between them is relatively small. One hundred years ago there was no such thing as a "detergent". Today, a "detergent" is anything that grabs onto dirt, soil, grease, oil, odor, bacteria, etc. and holds it well enough to loosen its grip on clothing, skin, hair, wall paint, etc. and pull it into the wash water. If the detergent hooks-up too well it may join the dirt on the fabric or surface and become part of the problem (soap and detergent scum). The cleaning agent must remain well attached to the water so that as the water rinses away, the detergent goes with it still holding fast to the dirts and soils it has released from the material being cleaned. Effectively, the cleaning agent allows insoluble particles to become soluble in water, so they can be washed away. This is how cleaning works whether it is called "soap" or "detergent". The category difference relates to how the cleaner (i.e., cleaning/washing agent) is created (i.e., materialized). A detergent or soap molecule is different on one end from the other. One end is attracted to "dirt" and the other is attracted to water.

Every washing agent contains chemicals to separate added material (a.k.a., "soil") from a surface (e.g., fabric) and carry it away with the water. These are called **surfactants (Read: surface active agents/ substances)** and they do whatever "cleaning" occurs. A properly designed and selected washing agent (i.e., surfactant) should remove matter efficiently from a surface and wash away completely without residue. In order to achieve this, a synergistic blend of two or three compatible surfactants is generally required.

Whereas soap is a metal salts of long chain higher fatty acids, detergents are sodium salts of long chain hydrocarbons like alkyl sulphates or alkyl benzene sulphonates. Soaps are prepared from vegetable oils and animal fats. Detergents are prepared from hydrocarbons of petroleum or coal and are "soapless". Soaps cannot be used effectively in hard water as they produce "soap scum" (i.e., insoluble precipitates of Ca2+, Mg2+, Fe2+ etc.). Detergents are generally not made insoluble by mineralized (so-called hard) water. The polar sulfonate (of detergents) is less likely than the polar carboxyl (of soap) to bind to calcium and other ions found in hard water. Some detergents also have secondary usages as acid solutions and foaming agents. Soap cannot be used as any form of acid solution, but it is sometimes used as a [component of a] lubricant (e.g., textile spinning). Whereas soaps are biodegradable, not all detergents are biodegradable.

- An example of a soap is sodium palmitate: $CH_3(CH_2)_{14}\text{-}COO^- Na^+$
- An example of a detergent is sodium lauryl sulfate: $CH_3(CH_2)_{12}\text{-}OS(O)_2\text{-}O^- Na^+$

Water, although a good general solvent, is also a substance with a very high surface tension. Because of this, water molecules generally prefer to stay together, rather than to migrate to other surfaces. **Surfactants** work by reducing the surface tension of water, allowing the water molecules to better wet the surface and thus increase water's ability to dissolve dirty, oily stains. Soap technology is essential for the removal of oil molecules from a surface. Oil molecules are non-polar, which means that they are not charged, and therefore, are not attracted to polar substances such as water (lipophilic). Because of this, oil tends to stick with its own molecules or other non-polar substances. Water is a polar substance which is made up of one positive and one negative charge. With this, water dissolves salt easily because salt is made up of charged ions in which the positive charge will be attracted to the negative ions in water (hydrophilic). When an appropriate surfactant is applied to oil, the lipophilic parts of surfactant will attach itself to the non-polar molecules of the oil. When water is applied onto this surface with a sponge, the hydrophilic component will be attracted to the water molecules and is lifted from the surface, together with the oil. Also, oil and grease will stick onto plates and cutlery during cleaning, and no amount of water can completely remove it without significant temperature and pressure, unless a surfactant is present.

> **NOTE:** *The presence of surfactants affects the natural micro-organism ecology.*

A surfactant (once dissociated in water) consists of a non-polar hydrocarbon tail and a polar head. The non-polar hydrocarbon tail interacts with non-polar substances through dispersion forces, whilst the polar head interacts with polar substances (normally water) which forms dipole-dipole interactions and hydrogen bonds with water. The presence of soap, for example, in a mixture of oil/grease and water will create an emulsion upon agitation. The hydrocarbon tails of soap dissolves in the oil and promotes droplet formation. These droplets repel each other, as they have the same net negative charge. These individual droplets can be dissolved in water as they form ion-dipole interactions.

An **emulsion** is a type of dispersion in which two normally immiscible substances are stabilised by another substance, called an emulsifier. For example, olive oil and water will not dissolve in one another, as their intermolecular forces differ (like dissolves like). While we can agitate to form a suspension, it is temporary and the oil and water will eventually separate into distinct layers.

5.3.1 Soap

Soaps are cleaning agents that are usually made by chemically reacting alkali (e.g., sodium hydroxide) with naturally occurring fat or fatty acids in order to facilitate the removal of organic and in-organic matter from the surface of a material. The metal is often an alkali metal such sodium or potassium, or an alkaline earth metal, such as calcium or magnesium. The fatty acid is often of plant or animal origin. Essentially, soap is a result of combining fat (i.e., oil) with an alkaline solution, and the productive chemical reaction is called **saponification**. The reaction produces sodium salts of these fatty acids, which improve the cleaning process by making water better able to lift away fatty material from skin, hair, clothes, and many other surfaces.

Historically, soap was made by boiling animal fat and adding lye to supply charged ends for the oil molecules. Time and temperature cause various combinations to occur and each will interact differently with the soils and especially with the calcium and magnesium in the wash water.

- Soap is most effective when mixed with mechanical pressure.
- Different oils and alkali solutions produce soaps with different properties. Especially coconut oil and palm oil produce excellent lathering properties.
- Temperature is a significant variable when using soap.
- Essential oils are sometimes added to soaps for their aromatic and anti-pathogenic properties.

5.3.2 Natural cleaners and soap

Natural ecologies produce a range of "soft" soap materials that are useful for cleaning. Although rare, soap is indeed found in nature. One family of historical soaps was made by infusing, simmering or mashing 'saponin' rich plants in water. If you have seen frothy puddles on the road near chestnut trees you have got the idea. Plants produce saponins as part of their immune system to deter insect attack and to act as natural anti-microbials, protecting their life bearing seeds. Hunter gatherers still exploit the soap-bearing plants for their medicinal properties and for cleaning. Grated or pounded horse chestnuts (aesculus hippocastanum), soaked and boiled bracken root (pteridium aquilinum), fern root (dryopteris filix-mas), snowberries (symphoricarpos albus), and soapwort (saponaria officinalis) will yield a soft water-based soap. Chestnuts are difficult to process, especially when you need to remove the shell first. Fern or bracken roots can not be collected from the wild and snowberries are best left on the plant as vital winter food for birds. Soapwort, our best known soap bearing plant is certainly worthy of its name and the whole of the plant can be used.

The soap nut tree Sapindus Mukorossi (a.k.a., Indian Soapberry) is a very large tree that produces prodigious amounts of a saponifying nut that you can use as a greywater safe laundry detergent, dish and hand soap. Soap nuts are berries that grow on a tree and naturally contain soap. Sapindus Mukorossi requires a fertile soil and a frost free climate. It's a tall tree that can take as long as ten years to begin fruiting. Just 5 berries can do multiple loads of laundry in a conventional washing machine. The natural soap found in these berries is called saponin. Saponin is a natural cleaner that works as a surfactant, breaking the surface tension of the water to penetrate the fibers of your clothing, lifting stains from the fabric, and leaving dirt suspended in the water that is rinsed away.

In an emergency when soap is not available, then there are many natural alternatives including: clean ash; vinegar, salt, mud, and sand/soil. Mud is a lesser known cleaner for hair (i.e., "shampoo") commonly used in the Mashreq and Maghreb regions.

Vinegar is a natural all-purpose acidic cleaner with anti-microbial properties. It is sometimes added to a surfactant solution to increase its effectiveness. Vinegar can be used to "clean" most standard surfaces including ceramic, glass, and fabrics. Vinegar cannot be used on marble or similar surfaces because its acid content will damage (i.e., "eat away") the surface, and quickly lose its shine.

Both baking soda and salt are excellent scouring powders. Baking soda is an excellent stain and odor remover. The solution can be used as a facial scrub. It can also be used as a toothpaste and tooth whitener. Clay can also be added to toothpaste to fa-ciliate whiter teeth due to its abrasiveness.

There are certain leaves that when diluted in hot water, form a weak cleaning solution. Two of the most well-known leaves for this purpose are palm an bay leaves.

5.3.3 Hydrophilic-lipophilic balance (HLB)

A system was developed to assist in making systemic decisions about the amount and types of surfactants needed in stable products. **Surfactants** are compounds that lower the surface tension (or interfacial tension) between two liquids or between a liquid and a solid. Surfactants may act as detergents, wetting agents, emulsifiers, foaming agents, and dispersants. The hydrophilic-lipophilic balance of a surfactant is a measure of the degree to which it is hydrophilic or lipophilic, determined by calculating values for the different regions of the molecule. The system is called the HLB system and has an arbitrary scale of 1-18. HLB numbers are experiment ally determined for the different emulsifiers. If an emulsifier has a low HLB number, there are a low number of hydrophilic groups on the molecule and it will have more of a lipophilic character. For instance, substances with low HLB numbers are generally oil soluble. As a result of their oil soluble character, they will cause the oil phase to predominate and form a water-in-oil emulsion. The higher HLB numbers would indicate that the emulsion has a large number of hydrophilic groups on the molecule and therefore should be more hydrophilic in character. Substances with high HLB numbers are water-soluble. And because of their water soluble character, they will cause the water phase to predominate and form an oil-in-water emulsion.

5.3.4 Detergent washing agent residue

Most commercial cleaners are filled with additional chemicals. Industry is not overly concerned with sustainable and efficient cleaning, and so, they create a chemical soup of additives. There are thousands of synthetic fragrances and over 200 brighteners which might be called dyes, colorants, color enhancers, color safe bleach, fluorescent whitening agents (FWA), etc. Then, manufactures add oils, silicones, and polymers to attach the fragrances and brighteners to the washing agents. Lubricants are added to make the fabrics less stiff, other lubricants for the washing machine pump, polyvinylpyrrolidone to seal the surfaces and ends of fibers, antiredeposition agents, perhaps sanitizers (anti-microbial agents), enzymes and oxidizers, also softeners, and of coarse inert fillers so you know by the size of the expensive containers that you got your moneys worth.

> **NOTE:** *In commercial products, the highest levels of fragrance are often found in "Baby" detergents. These fragrances can mask natural body odor/pheromones.*

Fabric softeners coat the surface of a fabric with chemical compounds that are electrically charged, causing threads to "stand up" from the surface, and thereby, causing the fabric to feel softer.

Oxidizers (oxidizing agents) and enzymes react with the smooth surface of fibers causing tiny pits and fissures that hold particles of residue. As temperature and humidity change, these imperfections swell and shrink. The eroded surface scatters light causing colours to look bleached and faded and the fibers weaken, loose memory, and eventually break. Elastomers fail significantly more quickly under these conditions than other fibers.

Optical brighteners are added to make the clothes appear cleaner. It would seem that if you can't clean clothes with less than six chemicals, maybe you're not focused on cleaning. Most of the other ingredients in commercial products are there to help the surfactants, or they are present to make you believe your clothes are clean, and hence, must remain in your clothes to be effective (e.g., optical brighteners).

Gutting (The Detergent Conspiracy) clarifies the effects of some modern detergents on clothing:

> *Tests done at the Clemson University School of Textiles and Polymer Science (Nicholson, 1995) showed that washing in regular grocery store detergent actually added a measurable amount of weight (contamination) to the clothes. Washing added 2% of the weight of the cloth in just 10 washings. The residue was equal to the full amount of detergent recommended to wash clothes. Let me be specific about this. When you pour in the detergent before the wash cycle, the scoop of powder or cup of liquid, you use is equal to the amount of chemical you will have in your clothes after 10 washes. It doesn't increase much beyond 10 washes because you reach a point where you are washing out as much as you are washing in.*
>
> *Detergents for colors and blacks have higher levels of polymer, which attaches to the clothing, to hide the scattering of light (i.e., "fading") caused by fiber erosion.*
>
> *The latest news in home laundry detergents is a Proprietary Protective Fiber Complex. This goes beyond polymers to a new silicone complex that not only provides additional lubrication, but also forms little cells of oiliness to keep the silicone complex and other care ingredients from washing away. This technology is especially important for high efficiency (HE) washing machines. The new machines not only save water and power, but rinse clothes more completely. Better rinsing is good for you but makes it difficult for the detergent makers to leave enough residue to assure you that your clothes are clean. Detergent makers had initially been forced to add more of the "care" chemicals, but this new technology allows them to overcome the great rinsing performance of your new HE machine and attach exactly the right amount of residue to your clothes.*

Of course, residue is not without consequence. It impacts the functioning of a material. It can degrade wicking, breathability, rapid drying, and water repellency in just a few washings. Detergent residue may not only slow drying, but will rehydrate itself when dry by drawing moisture from the air or from its user. Extra moisture reduces the efficiency of insulation. Insulation is also less able to loft up when coated with sticky residue, this further decreases effectiveness. Fabrics feel stiffer when coated with sticky chemicals because fibers cannot slip against each to allow a soft hand. Detergent manufacturers try to overcome some of this friction with silicone but deliberately leave some stiffness to ensure an opportunity to sell fabric softener. It is only a small exaggeration to say that washing clothes in regular detergents destroys them as fast as not washing them at all.

What are the effects of detergent residue on humans? An increasing slice of the population is experiencing episodes of eczema, psoriasis, and other reactive skin irritations loosely referred to as contact dermatitis. Humans will inevitably get the residue on their hands and likely end up consuming some of it. The human organ known as the skin may absorb some of it.

5.3.5 Overuse of soap

Cleanliness has come to mean an effort to remove germs and other hazardous materials. Today we now know that bacterial populations play an important role in the sustainment of a healthy ecological system, and in humans, a healthy immune system. Although cleanliness as the removal of dead cells, decomposition matter, contaminants, and excessive build-up of microorganisms and viruses has an important role in healthy functioning, excessive cleaning is often unnecessary and can damage surfaces.

It is the body oils forming a thin film on the surface of the skin that give human skin its smoothness. When the oils are washed off the body (i.e., stripped off of the skin) it tends to reduce the smoothness and "glow" of the skin. Now when you lather your body with soap during your bath, the lipophilic part attaches itself to your body oils and the hydrophilic part attaches itself to water. When you wash off the soap with water, your body oils are wrenched off from the skin and are washed of with the water. Your skin is deprived of its natural moisture but it gives you a feeling that you are cleaned/scrubbed. This feeling basically comes from the skin reporting the lack of an essential ingredient which we mistakenly feel to be fresh. You feel that you are cleaned off of all the "dirt" and that you are fresh. What really happened is that your body moisture is washed off and you need to replenish the body moisture. The soap industry realized what their soaps are doing to people and started adding moisturiser to the soap itself so that what is washed off by the soap part is replenished by the moisturiser. Excessive washing to remove oil from the skin also prevents vitamin D both from being formed on the skin and absorbed via the oils through the skin [via exposure to sunlight]. It is also relevant to note that soap is not essential for the removal of many pathogens from the hands.

Scholarly references

- Nicholson, R. (1995). *Detergent Evaluation of Sport-Wash*. Clemson University School of Textiles and Polymer Science. [atsko.com]

Book references

- Spellman, F.R. (2014). *The science of water: concepts and applications*. CRC Press.

Online references

- Guttin, D. *The Detergent Conspiracy*. Atsko. Accessed: January 7, 2020. [atsko.com]

TABLES

Table 12. Life Support > Water: *Drainage basin components.*

Drainage basin components	
Drainage basin component	**Provisional definition**
River channel	Linear feature along which surface water may flow, usually clearly differentiated from the adjacent flood plain or valley floor.
River reach	A homogeneous section of a river channel along which the controlling factors do not change significantly.
Channel pattern	Or channel planform, is the plan of the river channel from the air; may be either single threat or mult-ithread, varying according to discharge.
Floodplain	Valley floor area adjacent to the river channel.
Drainage network	Network of stream and river channels within a specific basin; may be perennial, intermittent or ephemeral/
Drainage basin or catchment	Delimited by a topographic divide or watershed as the land area which collects all the surface runoff flowing in a network of channels to exit at a particular point on a river.

TABLES

Life Support: Power Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

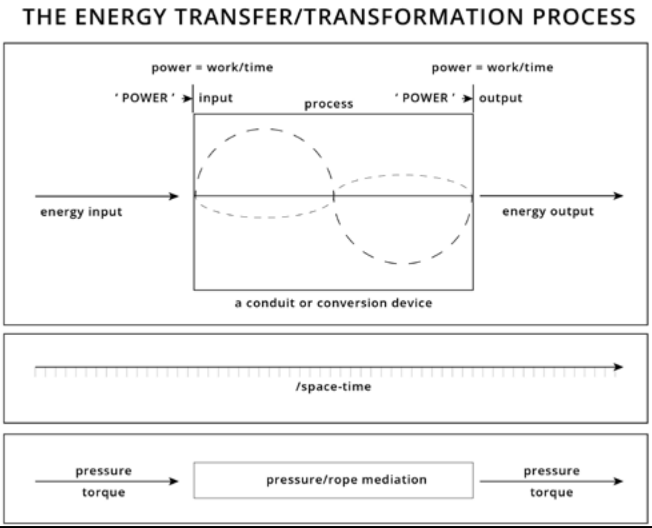
Keywords: power, power system, energy, energy system,

Abstract

A power service provides power. In physics, 'energy' and 'power' are extremely complex and convoluted concepts, which are themselves not fully understood. Hence, the practical view/perspective on 'energy' and 'power' taken by the Power Service System may be referred to as the substance-like view (or depository view) of energy and power. This view states that 'objects' move and 'energy' is a "substance-like" quantity that some objects/systems "have" and other objects/systems "need" in the "form of" a substance called 'power' in order to carry out functions. It is important to recognize that this is not how 'energy' and 'power' are defined in physics, but for practical application in a material system, this definition is appropriate at the service level. All power, currently, is motion power generation; the trading of one motion for another (i.e., all power generation, currently, relies on motion generation of a new system from a prior source of motion. The primary catalyst in the universe for energy production is hydrogen; the primary source in the universe for energy production is solar bodies. Power is always distributed through objects, because it is always objects that can only be in motion. Power can be distributed through tangible objects as conduits or guides of that power. Certain technologies require guided power; other technologies can use wireless power that moves electromagnetic objects in relation to one another. Many technologies require power to operate. The human life requirement of controlling body temperature can be extended through clothing (structure), but even more greatly extended through the powering of structures in such a way that the effects is an intentional change of temperature (e.g., a useful thermal output). Power's ability to change temperature (mass vibration) can be used to cook, provide warmth, process information, and display calculations. Energy can be stored, and the usage requirements for power can be calculated. Power generation and distribution technologies can be compared.

Graphical Abstract

Figure 19. *A power system transfers prior motion to another location and/or for another function through some other object that acts as a conduit or conversion device. Through this method, electricity, and other sources of power, can be produced.*



1 Energy (Introduction)

It is important to realize that in physics today, we have no knowledge of what energy is.
- Richard Feynman, Nobel Prize 1965, Physics Lectures 4-1

Although it is recognize that 'energy' is not a physical substance, from the perspective of the Habitat Service System, the substance metaphor shall be used. Herein, five principles guide the conceptualization and application of 'energy':

1. Energy can be viewed as *a substance-like, measurable quantity* [relative to all existence] that can be (1) stored in, (2) carried by, or (3) used over a specific time period by, an existent system.
2. Energy can be *transferred and flow* from one system (or carrier) to another, and by doing so, *effect change*. Energy provides the ability to run pre-set processes in systems (i.e., to "do things"). Energy is the instantiating capacity to produce effect\change.
3. Energy *maintains its identity* after being transferred.
4. The total quantity of energy is *conserved*.
5. The term 'energy' can be qualified by the name given to the system that is storing, carrying, or using the energy (e.g., kinetic energy, potential energy, internal energy, electrical energy, mechanical energy, etc.). Energy can exist in numerous carriers ("forms").

Energy is the capacity or ability to produce an effect (i.e., a change). The presence of a quantity of "energy" makes change possible. In physics, **energy** is a measurable property of existence (systems and objects) that can be transferred within and between existence, and converted/transformed into different forms by existence, but cannot be created or destroyed. In terms of a system, energy is the fundamental ability\capacity that allows for the existence\occurrence of the system, and any process, operation, and/or function performed by the system. It is an axiomatic property (and input) for the occurrence of any effect\change (and hence, all existence). In other words, energy is the capacity to produce change and sustain existence.

Energy is present if relationships between objects and fields change (in some way) from moment to moment, in time/memory. In physics, a 'field' is a physical quantity that has a value for each point in space and time. Moving things and doing actions in the physical environment requires energy.

CLARIFICATION: *A 'property' is any characteristic of a system. In physical systems, properties are considered to be either intensive or extensive. 'Intensive properties' are those that are independent of the mass of a system (e.g., temperature, pressure, density). 'Extensive properties' are those values that depend on the size or extent of a system. Therein, 'specific properties' are 'extensive properties' per unit mass.*

Hence, there exists for every system a property called 'energy'. As a property, energy is something that existence (e.g., matter) has, not something it is made of. In other words, energy is something possessed by a system or object, a property [of an something's existence relative to all existence] — a property of systems and objects that characterizes their existence, as well as their behavior and their relationships (interactions) to one another.

2 Power (Introduction)

Although it is recognized that power is the rate of transfer of energy, from the perspective of the Habitat Service System, the substance metaphor shall continue to be used. Herein, six principles guide the conceptualization and application of 'power'.

1. Power can be *generated* by the conversion of a source of energy.
2. Power can be *transmitted* and *distributed* from one system to another given a conducive pathway.
3. Power can be *modulated* given a control system that adjusts the quantity of energy and/or time-transfer rate the energy.
4. Power can be *applied* to (or *utilized* by) a conducive system to cause processes therein to operate.
5. The total quantity of power is *not conserved*.
6. The term 'power' can be qualified by the name of the system that is generating power, distributing power, or having power applied to it (e.g., electrical power, fluid power, hydroelectric power, wind power, etc.).

NOTE: *In the market, the electric power industry does not generate energy. They use the energy available in our common environment to produce and sell a product/service, that of, power as the controlled flow of electric charge (i.e., electricity or electric [potential] pressure). Through the movement of electric charge energy is transferred through useful circuitry that works/functions to provide us with additional (higher order[ed]) services.*

2.1 Energy and living systems

INSIGHT: *The more energy and power a system has [accessible], the more it can do. With more energy and more power there is more capacity/ability to do [new] things.*

"Energy" is necessary for existence; for the existence of movement, heat, electricity, and life. All organisms need energy to live. In other words, living organisms require available energy to stay alive (survive) and to thrive. Energy is connected to all organismal activities -- whenever organisms think or move, they "use" the energy available to their bodies.

NOTE: *Energy is a need (a critical requirement\ input) for the sustainment of an existent system.*

Organisms gathering together in the form of a 'city' have a need for energy [to maintain themselves and their material service systems]. Note that the processes of Earth's climate and ecosystem are driven by the energy that Earth receives from the sun, and the geothermal energy contained within the earth. The Sun is the main source of energy for the Earth, and in particular, for changes on the earth's surface. The sun is the major source of energy for phenomena on the earth's surface, including the growth of organisms, wind, ocean currents, and the water cycle. The Earth continuously receives an uninterrupted flow of energy from the Sun. For Earth's ecosystems, the major source of energy is [in the form of] sunlight. Note here that geothermal and nuclear energy do not derive from the subsequent transformations of solar energy, but are instead related to the Earth's formation.

HISTORICAL NOTE: *The first technological source of energy used by humans was fire, which hominids began using/controlling [at least] several hundred thousand years ago. Whereupon, they began using it to process/transform biomass, and then later, metals.*

Without energy there is no ability to maintain structure against the entropic movement of the universe -- without energy there is no ability to do anything [constructive]. Life is universally understood to require a source of energy and mechanisms [of action] (i.e., forces) with which to transform it. Energy in many of its forms may be used in natural processes, or to provide a service to society, such as heating, refrigeration, light, or performing mechanical work. Energy, in many of its forms, may be used in our bodies' own natural processes to keep our organisms functioning optimally, and we also use it to provide technological service to ourselves and our ecology through 'energy transformation', examples of which include heating, refrigeration, lighting, performing mechanical work to operate machines, and information processing. We can transform energy, and our transformations may have more or less potential [to facilitate the expression of our highest potential]. Human technology is based largely on the knowledge of methods to manipulate these "energy" forms to produce a desired function/outcome.

NOTE: *Technology channels energy into work or heat (i.e., service) for the function of human fulfillment, but it does not replace it.*

The Habitat Service System's Energy Sub-system takes energy [expressed as charge and/or motion] and redirects it (via transfer or transformation) into the iterative re-construction of our material environment for service continuity and higher functioning. In other words, the Habitat Service System redirects environmental energy into material constructions that facilitate our survival and the expanded emergence of our highest potential selves.

NOTE: *In physics, 'charge', also known as electric charge, electrical charge, or electrostatic charge (dielectric), and symbolized 'q', is a characteristic of a unit of matter that expresses the extent to which it has "more (-ion) or fewer (+ion) electrons than protons".*

At a biological level, energy transformation toward higher functioning occurs through a system known as 'mitochondria'. Biologically speaking, our organisms strip electric charge off food in a similar, though significantly more complex, manner to the functioning of turbines in a hydroelectric dam through which water flows. When water goes through the turbines electric charges are moved (by the technological system), and fed into an electric energy grid through electrical transmission. In fact, our mitochondria are a miniature example of this electric charge transport chain seen in hydroelectric or nuclear generating power systems. Living organisms take high energy [macro]nutrients as proteins, lipids, and carbohydrate (as well as solar energy) and process them through an energy transducer (i.e., "powerhouse") know as mitochondria to produce a set of "high" energy intermediaries (e.g., atp, nadph) that are then directed and delivered to regions of the cell(s) that maintain living function(e.g., muscles to contract or neurons to fire or digestive juices to be released or cells to replicate). This process is technically known as 'mitochondrial bioenergetics'. At a practical level, diet and lifestyle play an important role in sustaining the ability of mitochondria to transform energy for continued functioning of the body (and optimal health)

> **NOTE:** *A socio-economic system that increases available energy with equal access is 'egalitarian', and creates an environment where individuals have more freedom.*

2.2 The energy system architecture

> **CLARIFICATION:** *Energy transfer is required for the creation and operation of all existent systems.*

The Energy-Power System accounts for energy and directs power throughout the Habitat Service System. As a service system itself, the Energy-Power System harvests, stores, and transfers energy in order to generate and transmit power throughout the Habitat Service System in order to run all technical processes. 'Energy' and its transfer rate as 'power' are necessary for all technical change: the control of material resources, the [re] ordering of material existence, and the processing of all information. Without a source of 'energy', an its transfer rate as 'power', there is no ability to effect change in the world. Thus, the Energy-Power System inputs energy carrying resources (i.e., stores of energy) and outputs generated power, which is transmitted throughout the Habitat Service System.

> **NOTE:** *Every system in our physical environment needs energy to support its initial setup and sustained existence.*

The Habitat Service System Energy-Power Sub-system consists of:

1. Sources of energy.
2. An energy transfer (carrier conversion) system to generate power.
3. A heat rejection/thermal management system.
4. A power management and distribution system that includes controls for generation, transmission, and modulation of power.

The choice of a particular power system, and its particular architecture, will be determined by application requirements.

> **CONSIDER THE FOLLOWING QUESTIONS:**
> <u>*As a substance-like entity possessed by a system:*</u>
> *Where is the energy stored?*
> *Where did it come from?*
> *Where did it go?*
> *What does it do?*
> <u>*As the need of a system:*</u>
> *Into what system is the power generated?*
> *Through what medium is the power transmitted?*
> *Does the power meet required modulation parameters?*
> *What is the power use of the system?*

All organisms require a transfer of energy to live, and all real world systems require a transfer of energy to setup and operate. In other words, living organisms and their service systems require available energy to stay alive (survive) and to thrive. Energy is connected to all activities and changes in the real world -- whenever something occurs, energy is present (or transferred).

> **NOTE:** *Energy is a useful accounting tool that allows for calculating power sources and requirements, and determining whether change is "energetically" possible.*

Organisms gathering together in the form of a 'city' have a need for [the transfer of] energy to maintain themselves and their material service systems. Note that the processes of Earth's climate and ecosystem are driven by the [transfer of] energy that Earth receives from the sun, and the geothermal [transfer of] energy contained within the earth. The sun is the major source of energy for phenomena on the earth's surface, including the growth of organisms, wind, ocean currents, and the water cycle. The Earth continuously receives an uninterrupted flow of energy from the Sun. Note here that energy from geothermal and nuclear sources do not derive from the subsequent transfer of solar energy, but are instead related to the Earth's formation.

> **HISTORICAL NOTE:** *The first technological source of energy "used" by humans was fire, which hominids began "using" and "controlling" [at least] several hundred thousand years ago. Whereupon, they began "using" it to process/transform biomass, and then later, metals.*

Without energy there is no ability to maintain structure against the entropic movement of the universe -- without energy there is no ability to do anything [constructive]. Life is universally understood to require a source of energy and mechanisms [of action] with which to transfer (i.e., "transform") it. Energy in many of its carrying forms may be "used" in natural processes, or to provide a service to society, such as heating (temperature regulation), refrigeration, illumination, mechanization, and computation. Additionally, energy may be "used" by our bodies' own natural processes to keep our organisms functioning [optimally]. We can transfer energy, and our transfers of energy into and through material systems may provide us more or less potential to facilitate the expression of our highest potential.

> *"There is no energy in matter other than that received from the environment." – Nikola Tesla*

Human technology is based largely on the knowledge of methods of transferring energy into the generation of power for a desired service function and outcome/output.

> **INSIGHT:** *Energy-power technologies may be designed to channel energy into service (e.g., working and heating) for the function of human fulfillment, but it does not replace human fulfillment.*

The Habitat Service System's Energy-Power Sub-system takes energy and redirects it via technology into the operation and iterative re-construction of the Community's material environment for service continuity and higher functioning. In other words, the Habitat Service System redirects environmental energy into material constructions that facilitate our survival and the expanded emergence of our highest potential selves.

> **INSIGHT:** *Energy is [that which is required for] the temporal [re-]ordering of existence. A constant energy source and transfer as power is needed for maintaining the ordered state of living processes. An energy source is required fro controlling the ordered state of living systems.*

At a biological level, energy transfer toward higher functioning occurs through a system known as 'mitochondria'. Biologically speaking, organisms strip electric charge off food in a similar, though significantly more complex, manner to the functioning of turbines in a hydroelectric dam through which water flows. When water goes through the turbine-generator, electric charges are moved, and transferred into an electrical energy grid through electrical transmission. In fact, mitochondria are a miniature example of this electric charge transport chain seen in hydroelectric and nuclear generating power systems. Living organisms take high energy [macro]nutrients as proteins, lipids, and carbohydrate (as well as solar carrying energy), and process them through an energy "transducer" (i.e., a "powerhouse") know as mitochondria to produce a set of "high energy" intermediaries (e.g., atp, nadph) that are then directed and delivered to regions of the cell(s) that maintain living function (e.g., muscles to contract or neurons to fire or digestive juices to be released or cells to replicate). This process is technically known as 'mitochondrial bioenergetics'. At a practical level, diet and lifestyle play an important role in sustaining the ability of mitochondria to transfer energy for continued functioning of the body (and optimal health).

> **NOTE:** *A socio-economic system that increases available energy with equal access is 'egalitarian', and creates an environment where individuals have more freedom.*

2.3 Energy/power grid-network

> **INSIGHT:** *Possibly, energy is matter in motion relative to the rest of the matter in the universe.*

Within the infrastructure of the habitat service system, energy is transferred and power may be supplied via any of the following possible energy/power networks (grids):

1. Electrical power network (grid)
2. Mechanical power network (grid)
3. Pneumatic power network (grid)
4. Hydraulic power network (grid)
5. Gas transfer network (grid)

2.4 Energy-based services

> *"There is no energy in matter other than that received from the environment."*
> *– Nikola Tesla*

The Energy-Power System provides two main categories of service (two categories of output): heating services and non-heating services.

2.4.1 Heating services

Heating services are the services whose primary function is to deliver heat. Examples are space heating, water heating, and oven heating. These heating services have two important characteristics:

- An inherent inertia hence an inherent buffer storage capacity.
- The capability of being powered with a combination of heat and electricity, combination often very flexible.

The operation of heating services usually requires the simultaneous operation of less intensive, accessory non-heating services such as the operation of a pump or a control system.

2.4.2 Non-heating services

Non-heating services are all the other services that do not involve energy transfer via heat. It must be noted that the operation of these services may dissipate heat as a by-product, but it is not their prime function and more importantly it is not the amount of heat that will determine the level of operation of the service. These non-heating services have:

- No inherent inertia, hence no inherent buffer storage capacity.
- To be powered exclusively by electricity or [other] pressurized substances.

3 Energy carriers

Energy is a substance-like/information-like quantity that can flow or be transferred.

Per the substance-like perception of energy, energy is contained in and/or possess by what are called 'energy [re]sources' and 'energy carriers'. By definition, an energy source/carrier is a substance or a phenomenon that contains energy. In physics, energy always transfers (i.e., flows) simultaneously with at least one substance-like, physical quantity. Here, it is most appropriate to visualize energy as something that can flow from one place to another only when "carried" by another substance-like quantity through which change can be perceived. The thing for which change can be perceived and "carries" the energy is called an "energy carrier". Hence, the term "energy carrier" is able to provide clear language of how energy flows. The substance-like physical quantity which flows while energy is flowing, "carries" the energy, and may be referred to as an "energy carrier".

NOTE: *Transfer does not mean the same thing as transformation or conversion. The terms transformation and conversion mean that the thing itself changes form. Energy does not change form; it changes carrier or system.*

"Energy" is transferred between or within carriers. When energy is transferred, some change is occurring to the carrier(s), but no-thing (no transformation) is happening to the energy. And, the carrier(s) are not necessarily transformed or converted when the energy is transferred. The energy itself is not trans-formed or converted, because it doesn't have form; it is an abstraction (in physics) and remains energy regardless of the carrier or the value of 'energy' given to that carrier.

CLARIFICATION: *Energy cannot be transformed, it can only be transferred within and between carriers, until it is finally transferred through a service, whereupon it may be recoverable or irrecoverable.*

In this visualization, energy is not ever transformed (or converted) from one form into another, but rather, it transfers its carrier. The energy is transferred and the carrier(s) is changed (its motion and/or composition), possibly transformed/converted.

CLARIFICATION: *When people speak about "energy conversion", they mean converting one form of energy into another form. However, energy does not have different forms. There is just energy. Hence, "energy conversion" is a misconception.*

It is customary to say that energy exists in different forms, which are transformed or converted into one another during physical processes. However, using the term "energy form" for the respective categories is unsatisfactory because it easily leads to the misinterpretation that there are different kinds of energy. In other words, the notion that the energy is transformed leads to the incorrect idea that there are different forms of energy, which there are not. It is imprecise to speak about the forms of something that itself does not change, but rather, which only changes carriers. Energy maintains its identity regardless of transfer or material transformation. There is only one energy. In this sense, 'energy carriers' exist, whereas 'energy' is the result of a mathematical expression depicting the motion, action, or change of existence (of energy carriers).

> **NOTE:** *The utilization and generation of power (as the rate at which energy is transferred) always means the transfer of energy within and/or between carriers.*

Of course, there are limits as to how literally the expression "energy carrier" should be understood. The word "carry" implies only a temporal relationship between the flow of energy and the "flow" of an energy carrier. It is not meant to imply that energy and its carrier necessarily occupy the same position in space or even "flow" with the same velocity.

Further, an energy carrier can be "loaded" with more or less energy (as in, 'energy density') in the same sense that a carrier of material objects can be loaded with more or less of the objects.

The picture of "energy carriers" and "energy load factors" is especially useful to describe devices which are traditionally called "energy transformers" or "converters." Traditionally speaking, energy flows into an energy "transformer" in one form and out in another. Unfortunately, such language suggests that one physical quantity of energy is transformed into another within such a device. Instead, energy simply changes its carrier within the device. In other words, the energy is transferred from one carrier to another within the device. Accordingly, the term 'energy transceiver' is more appropriate to the actual function of such a device. A transceiver is a device composed of both a receiver and transmitter, and designed to transmit and receive energy (or data, or a signal).

> **NOTE:** *The common term for anything that is said to convert one form of energy into another (i.e., transfer energy between carriers) is a transducer. In common parlance, a "transducer" is anything that converts one form of energy into another. But, remember that energy cannot ever be converted or transformed, it can only be transferred between carriers. A* **transducer** *is a technological device that transfers ("transforms") energy from one form to another (from trans- "across" + dùcere "to lead"). The process of transferring energy between carriers (i.e., converting one form of energy to another) is known as* **transduction***. Of note, transducers are used in electronic communications systems to convert signals of various physical forms to electronic signals, and vice versa. Examples of transducers include a battery (energy carried by chemical composition which may be transferred to an electrically conductive circuit); a hydro-electric dam (energy carried by falling water transferred to an electrically conductive circuit).*

It is easy to visually represent the energy transfer from one device or region of space to another through an energy flow diagram (or energy transfer diagram).

> **INSIGHT:** *Talking about energy transfer stresses the importance of thinking about energy as staying the same kind of thing, but going from place to place.*

Energy carriers can be acquired, transported, and used. In this sense, energy is like information. We say that it can be stored in books, on computer hard drives, external drives, and disks. Information can be transferred from place to place via cables or by wireless transmission techniques. Information can be read and applied. But, there is nothing substantial about the information itself; it cannot be touched and its mass cannot be measured; it is substance-like. Even though information is moved from place to place and stored in different ways, and received by different people who apply it, nothing about the information itself has changed.

Hence, from the information metaphor (or analogy), three principles are present:

1. Energy can be viewed as a substance-like quantity that can be stored in an existent system.
2. Energy can "flow" or be "transferred" from one system to another and so cause changes.
3. Energy maintains its identity after being transferred. Energy is always energy.

3.1 The energy carrier function pyramid

Given the availability of technology, certain carriers of energy are more or less functional than other carriers. The presence and control of carriers higher up the pyramid allows for a more flexible and thought responsive environment than those lower on the pyramid. Electricity (electrical energy/power), for instance, is highly functional, because it can be converted to mechanical or thermal energy, and also used in electronics for a variety of functions, including but not limited to communication, computing, and lighting. Carriers higher up the pyramid are more functional than those lower on the pyramid.

When attempting to "convert" a quantity of energy to a form that is higher on the energy usefulness pyramid, invariably a large amount will be degraded. When thermal energy is converted to mechanical or electrical energy, part of the thermal energy has to be expelled into the environment. This energy is considered "degraded". Degraded energy still exists but essentially can no longer be converted into mechanical or electrical energy. In other words, degraded energy can no longer do work.

3.2 Energy carrying sources

> **NOTE:** *An energy carrying [re]source is something that can [be used to] produce heat, sustain organisms, move objects, or produce electricity.*

Generating and utilising power means transferring energy from one carrier/source to another. The transfer of energy, as well as generation and utilization of power, necessitates the acquisition and control of sources of energy (i.e., prime moving energy carriers). An **energy carrier** is a substance or sometimes a phenomenon that contains a quantity energy that can be transferred, and in doing so, produce work or heat, or to operate chemical or physical processes. An energy carrier does not produce energy; it simply contains energy imbued by another system. Energy carriers are the source of power for the Habitat Service System's energy-requiring systems. An originating source carrier of energy, prior to any processing by the Habitat Service System, is known as a 'primary energy source/carrier' (a.k.a., prime mover). This source/carrier may be used directly, or the energy therein may be transferred to a secondary (or tertiary) energy carrier prior to being transferred as power through useful service. Secondary and tertiary carriers occupy an intermediate step in the energy-supply chain between primary sources/carriers and end-use applications.

The term '**final energy carrier**' (a.k.a., "useful energy carrier") refers to the energy carrier that delivers the energy through intended end-service (i.e., end use). A 'final energy carrier' may be a primary, secondary, or tertiary energy carrier depending upon the number of intermediary transfers.

The concept of primary and secondary energy is used especially in energy statistics in the course of compilation of energy balances To avoid double counting, it is important to be able to separate new energy entering the system (primary) and the energy that is transformed within the system (secondary).

> **CLARIFICATION:** *'Fuels' are sometimes specifically and solely referred to as 'energy carriers'.*

There are two general types of energy source/carrier:

1. **Primary energy** is the state/source/carrier in which energy occurs in nature.
2. **Secondary energy** is produced by technically converting energy between carriers (forms). For instance, it can derive from primary energy through a single conversion step (solar radiation to electricity in a PV-panel) or through multiple steps from other forms of secondary energy (hydrogen from electrical energy through electrolysis). This conversion comes always with energy losses.

3.2.1 Energy carrier/resource development

> **NOTE:** *Humans can make devices that interact with the source of energy of all existence, and by so doing, increase their potential for creation and fulfillment.*

Energy development (a.k.a., energy resource development) is a field of scientific discovery and engineering focused on making available sufficient primary and secondary energy sources to meet power requirements.

3.3 Primary energy sources/carriers

> **NOTE:** *Energy in the universe transfers naturally over time in the presence of a triggering mechanism.*

A **primary source of energy (primary energy carrier)** is an energy carrying [re]source found in nature as an object or phenomena that has not been subject to any technical conversion or transfer process (by humans). It is a carrier in its "raw" form, and received as input into the Energy-Power System. The term 'primary energy' only designates those sources/carriers that involve extraction or capture, with or without separation from contiguous material, cleaning or grading, before the energy embodied in that source can be converted into heat or power.

In some cases, the primary energy carrier is the same as the final energy carrier (e.g., wood gathered for combustion and cooking purposes, animate power for pulling, or wind for sailing). Hence, primary carriers of energy can be used directly, such as burning wood sticks biomass for heat and light, or converted into a secondary energy carrying resource for storage/transport and/or higher functioning, such as wood pellets for a wood pellet stove.

The primary energy sources/carriers known to humankind are:

1. Biomass - organism composition (chemical motion)
2. Animate - organism motion (biochemical motion)
3. Solar - light/electromagnetic motion
4. Water (hydro) - type of planetary motion
5. Wind - type of planetary motion
6. Geothermal - type of planetary motion (heat; electromagnetic motion)
7. Mineral fuels (e.g., uranium) - mineral composition (atomic motion)
8. Fossil fuels (hydrocarbons) - fossilized organisms as composition (chemical motion)
9. Gravity - considered a type planetary motion (possibly, electric or electromagnetic motion)

It is observed that these primary energy carrying sources are not the ultimate source of origin of the energy. For instance, animate comes from biomass, whereas biomass ultimately comes from the sun. Apart from geothermal and mineral fuels (a.k.a., "nuclear"), all "primary energy carrying sources" ultimately get their energy from the sun.

> **NOTE:** *The systems that transfer and/or convert this primary source energy are sometimes called 'primary energy conversion/transfer systems'.*

3.3.1 Fuels and flows

Besides gravity, there are two types of primary energy carrier: fuels and flows. Fuels like coal, natural gas, and uranium are dense carriers of energy, that are transformed/converted (i.e., "consumed") when used. Flows are natural [motion] processes that carry energy associated with their movement. Using a flow means harnessing the motion of that flow in order to transfer its carried energy.

Fuels are dispatchable, which means the energy is available for transfer whenever it is needed. A flowing carrier differs from a fuel, because energy transfer from a flowing carrier is only available when the carrier is flowing. For instance, the energy carried by solar radiation is only available when there is sunlight, and the energy carried by wind is only available when the atmosphere is flowing (i.e., when it is windy).

3.3.2 Renewability

A.k.a., Continuity.

Primary energy carrying sources may be classified according to their **renewability** (as in, re-new-able versus non-re-new-able). However, this terminology is rather ambiguous, as the meaning of the word "renewable" often depends on the context of its use. In general, a renewable energy carrier (i.e., "renewable energy") refers to "inexhaustible natural resources", and is contrasted with in-earth exhaustible natural resources (fossil fuels). Hence, energy carrying resources are considered 'renewable' if they are naturally replenished (in a relatively short time-frame).

Presently, there are seven known renewable energy carrying sources:

1. Biomass - fuel
2. Animate - flow
3. Solar (strictly intermittent) - flow
4. Hydro - flow
5. Wind (strictly intermittent) - flow
6. Geothermal - flow
7. Gravity - unknown

Hydrocarbons are primarily contained in coal, oil and natural gas (Read: in-earth hydrocarbons). Some plants also contain hydrocarbons, but these would be classified as biomass sources. Of note, some in-earth forms of hydrocarbon are actually "renewable" in terms of being naturally replenished, but they take a long time (in concern to a human lifespan) to renew. Additionally, it is possible to imagine that a species could draw, harvest, or transfer so much motion from its planet[ary motion cycles] through large scale geothermal or atmospheric wind collection that the draw on those sources could be unsustainable and disrupt the natural motions of the planet. One could also use biomass at such a rate that it too becomes unsustainable.

> **DEFINITION:** *A variable renewable energy (VRE) carrier is a renewable resource that is non-dispatchable due to its fluctuating nature, like wind and solar, as opposed to a controllable renewable sources such as hydroelectricity, or biomass, or a relatively constant source such as geothermal or run-of-the-river hydroelectricity.*

3.4 Secondary energy sources/carriers

A **secondary source of energy (secondary energy carrier)** is derived from the transfer of energy from a primary energy carrier, whereupon the carrier itself may or may not have been transformed/converted in the process. Secondary energy should be used to designate all sources of energy that results from transformation of primary sources. Secondary sources of energy are sometimes confusingly referred to as just "energy carriers", wherein primary sources are referred to as "energy sources", because secondary carriers, unlike primary sources that are not also 'final energy carriers', are generally capable of being stored in a usable "form" and transported in a controlled manner from one place to another.

> **CLARIFICATION:** *Primary carriers transfer energy directly from the environment, while secondary carriers acquire energy transferred from the primary environmental carriers.*

There are four types of secondary energy carrier:

1. Mechanical [solid] carrier - mechanical energy/power
2. Pressurized [fluid] carrier - fluid energy/power (including elevation relocated water or other liquid)
3. Chemical [bond] carrier - chemical energy/power
4. Electrical [charge] carrier - electrical energy/power or electricity
5. Electromagnetic [field] carrier - electromagnetic energy/power
6. Thermal carrier - thermal energy/power or internal energy

> **NOTE:** *The systems that transfer and/or convert this secondary source energy are sometimes*

called 'secondary energy conversion/transfer systems'.

For example, petrol fuel (secondary, chemical carrier) is made from the processing of crude oil (primary). Electricity (secondary) may be obtained from the harvesting of planetary motion (hydro-electric and wind-electric). Note that a battery is an example of a secondary energy source, a type of 'fuel' that stores electric charge potential as "chemical energy". Electricity is a secondary energy resource, and it can be generated/made by a number of different primary sources. What we commonly know as "the flow of electric charge/power" (i.e., "electricity") is [to us] a secondary "energy" source. The controlled flow of electric charge (i.e., "electricity") is a product of the transfer of energy from primary sources of energy such as wind, coal, natural gas, or solar [into the controlled flow of electric charge].

NOTE: *The electrical carrier takes the form of an electrical cable network or electrical grid reticulating electricity around the generation point. The thermal carrier may take the form of a pipe network reticulating heating "hot" water (HHW) around the system. Heat is then delivered to the services through heat exchangers. Operation of an HHW network requires a minimum of electricity for the circulating pumps and controls.*

The presence of thermal carriers is not strictly necessary in a habitat service system since all heating services could be powered exclusively by electricity (heat on-demand), but the presence of a thermal carrier is generally well justified by the facts that: 1) a large amount of waste heat can be recovered on the electricity generation process; and 2) heat-only can be produced much more efficiently than electricity-only.

The concept of 'renewability' does not apply to secondary energy sources/carriers. For instance, the energy sources we use to produce (make/generate) the controlled flow of electric charge may be characterized as renewable or non-renewable, but electric charge (and its flow) cannot be classified as either renewable nor non-renewable.

3.5 Power systems

NOTE: *Conceptually, there is no such thing as an "energy system", because energy is just a quantity. Power involves the transfer of energy per time, and hence, involves a set of relationships that form a system.*

All technological systems require the transfer of energy for their construction and operation. In order to transfer energy effectively, the flow of its carriers must be controlled. The rate at which energy is transferred is called power. Whereas an 'energy system' may be said to account for the presence of and necessity for energy, a 'power system' may be said to control the rate and quality of energy transfer, by controlling its generation, transmission, distribution, and modulation.

Power systems may be categorized according to the type of carrier experiencing the transfer of energy. Note that the suffixes of the types of power system mentioned below, end in either "ic" or "ical" or "al", which are used to form adjectives from nouns (gerunds) with the meaning "of or pertaining to" or "a type of". The suffix "ic" also means "application of", as in electron*ic* (the application of electrons) or atom*ic* (the application of the atom).

Energy can be transferred, and hence, work can be done in the following physical [power] systems:

1. **Kinetic/mechanical power system (mechanical power system)** - a solid is the carrier using linear or rotational motion.
2. **Fluidic/fluidal power system** - a fluid is the carrier.
3. **Atomic/chemical power system** - the structural composition of atoms and molecules (i.e., mass, number of particles, and bonding), is the carrier.
4. **Thermic/thermal power system** - atomic and molecular oscillation in all degrees of freedom is the carrier.
5. **Electric/electrical power system** - electrically conductive circuit within which free charged particles are the carrier.
6. **Magnetic power system** - magnetism is the carrier.
7. **Electromagnetic power system** - an electromagnetic radiating "wave" (the vacuum or ether) is the carrier.

NOTE: *Technically, mechanical power can be subdivided into solid mechanical, fluid, and inertia.*

These power systems may be connected to form a network of transceiving (transmitting and receiving, conversion) power systems.

NOTE: *As a physical concept, 'power' requires both a change in the universe and a specified time over which the change occurs.*

When building a system to transfer energy for the production of power, three main questions must be considered:

1. What is the original carrier of the energy?
2. What energy transfer process will be used?
3. How will the carrier be changed and/or moved from one place to another?
4. How will the energy eventually be transferred through useful service?
5. Other factors that must be considered include where the energy carrier is located, the amount of

power that must be produced, and the length of time it must be controlled.

Like all technology systems, power systems have inputs, processes, outputs, and feedback. All power systems require the same five resources as inputs:

1. Information
2. Materials
3. Tools and machines
4. Energy
5. Time
6. ~~Capital~~ // no market in community

The operation of a power system requires:

1. Power generation units
2. System controls
3. System stability

3.6 Energy transfer/power conversion systems

> **NOTE:** *Power engineering deals with the generation, transmission, and distribution of power as well as the design of a range of related devices.*

Power conversion refers to the time interval transfer of energy between different carriers/sources of energy (Note: In practice, "power conversion" is sometimes referred to as "energy conversion"). A power conversion system (a.k.a., "energy transfer system", "energy transformer", "energy transducer", "energy transceiver", "energy converter") accepts input energy as power from one carrier (i.e., one power/energy system) and delivers output energy as power to another carrier (i.e., in another energy system).

> **REMEMBER:** *In the substance-like metaphor, energy does not ever transform, though it can be transferred and stored. And, power (as the rate of transfer) does not convert, though it can be generated, distributed, and utilized.*

The 'utilization' and 'generation' of power always means transferring energy from one carrier to/through another. Whereas the 'generation' of power relates to a source of energy, the 'utilisation' of power serves an end-use of energy. In between, the energy can flow through a number of energy transfer/power conversion steps. The words "generation" and "utilisation" are a little confusing because, in fact, no energy can be created or destroyed, but power can be supplied and cancelled.

The generation, transmission, and utilization of power requires the input, transfer, and output of energy:

1. <u>When generating power</u>, energy is made available (input) from a source, and transferred into/through a technical system to produce power (energy transferring at a specific rate).
2. <u>When transmitting power</u>, energy is carried by the transmitting system at a specific rate.
3. <u>When utilising power</u>, energy is made unavailable (output), possibly irretrievably so, to power a process or service.

Power conversion devices are not 100% efficient. Some input energy is "lost" in the transfer process.

Energy transfer/power conversion devices (technologies and systems) are generally named for their input energy carrying system:

1. Mechanical power conversion transfers energy carried by a mechanical system to:
   - Mechanical > fluidal (e.g, power steering pump)
   - Mechanical > electrical (e.g., alternator)
   - Mechanical > fluidal (e.g., fan, propeller)
   - Mechanical > thermal (e.g., thermal welding)
2. Fluidal power conversion transfers energy carried by a fluidal system to:
   - Fluidal > mechanical as linear/rotational motion (e.g., turbine)
3. Electrical power conversion transfers energy carried by an electric system to:
   - Electrical > mechanical (e.g., electric motor, actuator)
   - Electrical > thermal (e.g., heater, light bulb)
   - Electrical > electromagnetic (e.g., antenna transmitter)
   - Electrical > magnetic (e.g., electromagnetic induction, electrical transformer)
4. Magnetic power conversion transfers energy carried by a magnetic system to:
   - Magnetic > electrical (e.g., electrical generator)
   - Magnetic > mechanical (e.g., magnets)
- Electromagnetic power conversion transfers energy carried by a electromagnetic system to:
   - Electromagnetic > electrical (e.g., antenna receiver, solar panel)
   - Electromagnetic > mechanical (e.g., electromagnet)
5. Thermal power conversion transfers energy carried by a thermal system to:
   - Thermal > fluidal (e.g., steam plant boilers generate electricity)
   - Thermal > mechanical (e.g., combustion engine)
   - Thermal > electric (e.g., thermopile, thermoelectric generator)
6. Chemical power conversion transfers energy carried by a chemical system to:
   - Chemical > mechanical (e.g., chemical motor, internal combustion engine)

- Chemical > electrical (e.g., chemical battery)
- Chemical > electromagnetic (e.g., combustion fire)

Power can be converted from one system to another form in three primary ways:

1. Through the action of forces.
   A. **Electric and magnetic [force] fields** - Charged "particles", upon which electrical fields exert forces, possess potential energy in the presence of an electric field in a way similar to that of an object in a gravitational field. These force fields can accelerate particles, converting a particle's potential energy into kinetic energy. Likewise, charged particles can interact via the electric and magnetic fields they create, transferring energy between them, and in the case of an electrical current in a conductor, cause molecules to vibrate (i.e. converting electrical potential energy into heat).
   B. **Frictional forces** - The macroscopic (large-scale) energy of an object, that is, the potential and kinetic energy associated with the position, orientation, or motion of the entire object, not counting the thermal or heat energy of the system, can be converted into thermal energy (heat), whenever the object slides against another object. The sliding causes the molecules on the surfaces of contact to interact via electromagnetic fields with one another and start vibrating.
   C. **Gravitational force** - when gravity accelerates a falling object it converts its potential energy to kinetic energy. Likewise, when an object is lifted, the object stores the energy exerted by the lifter as a potential energy in the earth-object system.
2. When atoms absorb or emit electromagnetic radiation. When light falls on an object, an incident photon may either pass through the object, be reflected by the object, or be absorbed by the atoms making up the object. If most of the photons pass through, the object is said to be transparent. Depending on the smoothness of the surface on the scale of the photon's wavelength, the reflection may be either diffuse (rough surface) or coherent (smooth surface). If the photon is absorbed, the photon's energy may also be split up and converted in the following ways:
   A. **Photothermal effect:** the energy absorbed may simply produce thermal energy, or heat in the object. In this case the photon's energy is converted into vibrations of the molecules called phonons, which is actually heat energy.
   B. **Photoelectric effect:** the energy absorbed may be converted into the kinetic energy of conduction electrons, and hence electrical energy.
   C. **Photochemical effect:** the energy may bring about chemical changes which effectively store the energy.
3. When nuclear reactions occur, that is, when there are rearrangements of the subatomic particles that make up the nuclei of atoms. There are two basic types: **Fission** - when nuclei combine, and **Fusion** - when nuclei split apart.

3.7 Transfer efficiency

> **INSIGHT:** *In community, all energy transfer loses are heat and/or technical losses; there are no administrative loses. In the market, many energy transfer losses are administrative losses.*

Efficiency' is the ability to achieve a desired result with as little "loss" of energy and effort as possible. In concern to energy and power, it is the ability to avoid "wasting" materials, energy, efforts, and time in producing a desired result. It is a dimensionless performance measure of a process or technology. The term 'efficiency' makes sense only in reference to the wanted effect. Energy transfer efficiency is not defined uniquely, but instead depends on (is relative to) the usefulness of the output. An incandescent light bulb, for example, might have 2% efficiency at emitting light, yet still be 98% efficient at heating a room. (in practice it is nearly 100% efficient at heating a room because the light energy will also be converted to heat eventually, apart from the small fraction that leaves through any windows).

> **NOTE:** *Energy may be transferred within or between carriers at various efficiencies, depending upon technical ability.*

Transfer efficiency refers to the ratio between the useful output of a system, and the input, and it can be calculated in terms of energy and power. Efficiency is directly calculated through the output-input ratio, where the output is the desired service, and the input is the quantity input into the system. In concern to energy, the efficiency of an energy transfer is the percentage or fraction, of the energy input that is transferred to useful output. This figure is multiplied by 100% to give you the result in percentage.

> **NOTE:** *The system boundary must be carefully specified when measuring efficiency.*

Generally, energy transfer efficiency is a dimensionless number between 0 and 1.0, or 0% to 100% (when the ratio is multiplied by 100). Transfer efficiency is usually expressed by the Greek letter η = (output energy / input energy) x 100%.

1. In concern to energy, the efficiency of a system transfer process is defined as the "quantity of

energy" output from the transfer (the output) divided by the "quantity of energy" put in for transfer (the input), and then, multiplied by 100%.
   - Energy efficiency % = (useful energy output / useful energy input) x 100%
   - Efficiency = useful energy out / total energy in
2. In concern to power, the efficiency of a system transfer process is defined as "useful power output" divided by the "total power consumed", and then, multiplied by 100%.
   - Power efficiency % = (useful power output / useful power input) x 100%
   - Efficiency = useful power out / total power in
3. In concern to work, efficiency is the ratio of useful work out from the total amount of work done, as a percentage.
   - Efficiency % = (useful work out (J) / Total work done (J)) x 100

If the efficiency of an energy transfer amounts to 60%, this means that out of 100 energy units included in a process (total energy in), 60 were transferred through to desired change (useful energy out), whereas the other 40 were transferred through undesired change (wasted energy out). That energy or power which has gone into the process, but has not come out of the process as useful is generally considered "wasted" or a loss (i.e., cannot be used).

> **NOTE:** *Energy is always transferred from one input carrier to many output carriers (the application + "losses") with perfect efficiency (law of conservation). An energy transfer technology with an efficiency over 100% may be called an 'overunity machine' or 'zero-point machine', and there are no known schematics for such a machine, which breaks the laws of thermodynamics.*

Electrical power transfer has three types of energy loss:

1. The Joule effect, where energy is lost as heat in the conductor (a copper wire, for example).
2. Magnetic losses, where energy dissipates into a magnetic field.
3. The dielectric effect, where energy is absorbed in the insulating material.

4 Power system types

There are seven general power system types:

1. Mechanical (kinetic) power systems
2. Fluidic/fluidal power system
3. Atomic/chemical power system
4. Thermic/thermal power system
5. Electric/electrical power system
6. Magnetic power system
7. Electromagnetic power system

4.1 Mechanical (kinetic) power systems

Mechanical power is the time rate of motion, and when motion is applied to a task, then it is the time rate of work [done to accomplish the task].

4.2 Mechanics

Mechanics is (physics) an area of science that studies and attempts to predict the behavior of physical bodies when subject to forces and displacements, and the subsequent effects of the bodies on their environment. In other words, it is concerned with the action of forces that displace (i.e., move) material objects with mass. It is the study of interactions between bodies and forces that produce motion. Mechanics describes the motion of bodies, and the causes that effect them. This includes the special case where the "motion" is no motion (i.e. bodies that are stationary).

> **NOTE:** *In physics, a physical body or physical object (sometimes simply called a body or object; also: concrete object) is an identifiable collection of matter, which may be more or less constrained by an identifiable boundary, to move together by translation or rotation, in 3-dimensional space. In classical mechanics a physical body is collection of matter having properties including: mass, velocity, momentum, and energy. The matter exists in a volume of three-dimensional space. This space is its extension. In continuum mechanics an object may be described as a collection of sub objects, down to an infinitesimal division, which interact with each other by forces which may be described internally by pressure and mechanical stress.*

Mechanics is based on the three "Newtonian laws of motion":

1. **First law** - In an inertial reference frame, an object either remains at rest or continues to move at a constant velocity, unless acted upon by a net force.
2. **Second law** - In an inertial reference frame, the sum of the forces F on an object is equal to the mass m of that object multiplied by the

acceleration a of the object: F = ma.

3. **Third law** - When one body exerts a force on a second body, the second body simultaneously exerts a force equal in magnitude and opposite in direction on the first body.

There are two branches of mechanics, each with multiple sub-branches:

- **Classical mechanics** - study of motion above atomic scale.
- **Quantum mechanics** - study of motion below atomic scale.

These branch distinctions, however, are not fundamental in nature. They are distinctions made by humans, which are useful for comprehending and for engineering when nature (i.e., physics) is not fully understood. Hence, any attempt to define the difference is to some extent arbitrary, and will not last as the subject is unified through greater understanding.

4.2.1 Energy in mechanics

In concern to mechanics, there are four elements that compose the characterization of energy:

1. **Cause** - that which allows for (gives rise to) a force.
2. **Force** - any interaction that, when unopposed, will change the motion of an object.
3. **Displacement (effect)** - a motion of any distance.
4. **Work** - the act of using a force to cause displacement. Hence, there are three key elements to the concept of 'work' - force, displacement, and cause.

Energy is what allows the exertion of a force. With "zero" energy, no force can be exerted, and no action can be taken. By exerting a force over a distance energy is transferred ("transformed"). Work is technically defined as what is done when a force moves its point of application. Work implies movement of a body by the application of a force. Work may be perceptible or imperceptible to human senses. Energy, can therefore be termed as that which can bring about a change and is the cause of all dynamic manifestations.

4.2.2 Classical mechanics

"Classical" mechanics is a branch of physics concerned with the set of physical laws describing the motion of matter (in macroscopic and microscopic form, but not atomic) under the influence of forces. Within classical mechanics are fields of study that describe the behavior of solids, liquids and gases and other specific sub-topics.

Classical mechanics consists of the work mostly done in the areas of chemistry and physics prior to the 20th century. This includes the organization of the periodic table, thermodynamics, the wave theory of light, and Newtonian mechanics.

There are three perspectives (branches) deriving axiomatic formulations for classical mechanics separated into two forms of mechanics, the original Newtonian Mechanics and the reformulated Analytical Mechanics:

Newtonian mechanics (original) is classical mechanics based on the Newtonian understanding of motion as understood through the equation (Newtonian second law of motion):

- Force (f) = mass (m) · acceleration (a)
- This formulation of classical mechanics is also widely known as Newtonian mechanics.

Analytical mechanics (reformulation of Newtonian mechanics with an emphasis on system energy, rather than on forces):

1. Classical mechanics based on the formulation of Lagrangian mechanics.
2. Classical mechanics based on the formulation of Hamiltonian mechanics. In Hamiltonian mechanics, a classical physical system is described by a set of canonical coordinates r = (q,p), where each component of the coordinate q_i,p_i is indexed to the frame of reference of the system. A particle on a line whose position (q) and momentum (p) are functions of time (t). If the energy (H) is a function of position and momentum, then the time evolution of the system is:
   - $dp/dt = -(\partial H/\partial q)$ and $dq/dt = +(\partial H/\partial p)$
   - where, H = H(q,p,t) is the Hamiltonian, which often corresponds to the total energy of the system.

NOTE: *In physics, a force is any interaction that, when unopposed, will change the motion of an object. In other words, a force can cause an object with mass to change its velocity (which includes to begin moving from a state of rest), i.e., to accelerate. Force can also be described by intuitive concepts such as a push or a pull. A force has both magnitude and direction, making it a vector quantity. It is measured in the SI unit of newtons and represented by the symbol F.*

The following branches of Classical Newtonian Mechanics are useful to have in awareness, simply because they are used so frequently in academic and industrial physics and engineering:

1. **Statics** is (mechanics) the study of forces in equilibrium without consideration of changes over time (stationary objects).
   - The study of equilibrium and its relation to forces.
2. **Kinematics** is (mechanics) the study of (relative) motion, including displacement, velocity, and acceleration, without any consideration of why

those quantities have the values they do. Herein, the description of the motion itself (expressed by mechanics) is called kinematics. In other words, kinematics is the branch of classical mechanics that studies and describes the motion of physically joined points/parts (multi-link systems) such as an engine, a robotic arm or the skeleton of the human body. These descriptions setup the relevant degrees of freedom, represented as variables in a relevant mathematical form. Kinematics is concerned with the effects of motion on objects without reference to its causes. Kinematics pertains to motions determined by conservation laws: kinematics tells you that momentum and energy have to balance.
  - The study of the implications of observed motions without regard for circumstances causing them.
3. **Kineto-statics** is (mechanics) is concerned with the study of forces in equilibrium, with the addition of motion related forces (like inertia forces via D'Alembert's principle) one instant at the time. Results from one time frame do not affect the results on the next time frame.
4. **Dynamics** is (mechanics) concerned with the effects of forces and torques on the motion of objects. Dynamics means a study of the rules governing the interactions of particles, which allow for a determination of why the quantities have the values they do. The description of the causes of motion expressed by mechanics, and how these causes effect motion is called 'dynamics'. These causes are often divided into forces and torques. Dynamics provides full consideration of time varying phenomena in the interaction between motions, forces and material properties. Typically there is a time-integration process where results from one time-frame effect the results on the next time-frame. Dynamics depends on interactions, and not just on conserved quantities, which is what kinematics depends on.
  - The study of motion and its relation to forces.
  - In engineering, **dynamics** is sometimes referred to as the combination of kinematics and kinetics of proper motion. And, statics is the kinematics and kinetics of static equilibrium.

The difference between kinematics and dynamics may be understood in terms of programming a computer to simulate the physical system. 'Kinematics' is the data structure required to simulate the general situation, involving what variables with what range of values. 'Dynamics' is the actual algorithm that simulates the motion. The difference is the consideration of forces. 'Kinematics' concerns the range of movement or change a system can undergo, or the state space in which it acts. 'Dynamics' concerns the movement it undergoes according to the laws of motion. This means that conservation of energy and other quantities is dynamical, because it only holds when the equations of motion are in effect.

When mechanics is studied in the context of macroscopic bodies and systems, then following branches are useful to have in awareness, simply because they are used in academic and industrial physics and engineering:

- Continuum mechanics is the study of the physics of continuous materials, and can be subdivided into solid mechanics and fluid mechanics. It is a branch of Classical Newtonian Mechanics that deals with analyses of the kinematics and mechanical behavior of materials modelled as a continuous mass, rather than as discrete particles.
  - **Solid mechanics** is the branch of continuum mechanics that studies the behavior of solid materials, especially their motion and deformation under the action of forces, temperature changes, phase changes, and other external or internal agents. The study of the physics of continuous materials with a defined rest shape.
  - **Fluid mechanics** is the branch of continuum mechanics that studies the behavior of fluids (liquids, gases, and plasmas), and the forces on them. The study of the physics of continuous materials which deform when subjected to a force.
  - **Deformation mechanics** is the branch of continuum mechanics that studies the behavior of a body undergoing transformation from a reference configuration to a current configuration.
  - **Rheology** is the branch of continuum mechanics that concerns the study of materials with both solid and fluid characteristics.

When mechanics is applied to the design and development of technology, then following branches/ disciplines are useful to have in awareness, simply because they are used in academic and industrial physics and engineering:

1. **Applied mechanics** is the practical application of mechanics.
2. **Mechanical engineering** (classical mechanical engineering) is a discipline of engineering that applies the principles of physics and materials science for analysis, design, manufacturing, and maintenance of mechanical systems. It is the branch of engineering that involves the production

and usage of heat and mechanical power for the design, production, and operation of machines and tools.

3. **Biomechanics** is (mechanics) concerned with the study of the movement of living things using the science of mechanics, which provides conceptual and mathematical tools as necessary for understanding how living things move. Biomechanics is the study of the structure and function of biological systems such as humans, animals, plants, organs, fungi, and cells by means of the methods of mechanics.
4. **Mechatronics** is multidisciplinary field of science that includes a combination of mechanical engineering, electronics, electrical engineering, and computer engineering (and possibly other disciplines) in the design and development of technology.

NOTE: *In a mechanical-electronic system torque is analogous to current, and speed is analogous to voltage. The product of speed and torque is power (mechanical) and the product of current and voltage is power (electrical).*

4.3 Thermodynamics

Simplistically, thermodynamics is a branch of physics that deals with heat and temperature and their relation to energy and work. In other words thermodynamics studies the effects of changes in temperature, pressure, and volume on physical systems on the macroscopic scale, and the transfer of energy as heat. The behavior of these quantities is governed by the four laws of thermodynamics, irrespective of the composition or specific properties of the material or system in question. The laws of thermodynamics are explained in terms of microscopic constituents by 'statistical mechanics'. The goal of 'statistical mechanics' is to extract the macroscopic thermodynamic quantities like pressure, entropy, internal energy, etc. in terms of the microscopic laws governing a particle.

Thermodynamics was established in the 19th century as scientists were first discovering how to build and operate steam engines. Thermodynamics deals only with the large scale response of a system which we can observe and measure in experiments. In more detail, the theory of the relations between various macroscopic observables, such as temperature, volume, pressure, magnetization, and polarization of a system is called 'thermodynamics'. There are two principal forms of thermodynamics as a study: classical and relativistic.

The starting point for most thermodynamic considerations is the laws of thermodynamics, which postulate that energy can be exchanged between physical systems as 'heat' and 'work'. They also postulate the existence of a quantity named entropy, which can be defined for any system. In thermodynamics, interactions between large ensembles of objects are studied and categorized. Central to this are the concepts of system and surrounding. Therein, a system is composed of particles whose average motions define its properties, which in turn are related to one another.

Classical thermodynamics describes how heat flows in order to maximize total entropy, and is valid only for systems at rest with respect to observer. Relativistic thermodynamics was primarily introduced to account for the effect of relative motion between the observer and the system.

In concern to classical thermodynamics, there are several initial "laws". The zeroth law of thermodynamics involves some simple definitions of 'thermodynamic equilibrium'. Thermodynamic equilibrium leads to the large scale definition of 'temperature', as opposed to the small scale definition related to the 'kinetic energy' of the molecules. The first law of thermodynamics relates the various forms of kinetic and potential energy in a system to the work that a system can perform and to the transfer of heat. This law is sometimes taken as the definition of 'internal energy', and introduces an additional state variable, 'enthalpy'. The first law of thermodynamics allows for many possible states of a system to exist. But, experience indicates that only certain states occur. This leads to the second law of thermodynamics and the definition of another state variable called 'entropy'. The second law stipulates that the total entropy of a system plus its environment can not decrease; it can remain constant for a reversible process but must always increase for an irreversible process.

The basic equation in thermodynamics is:

- $dE = dQ - PdY$
- Where, E is internal energy of a subsystem; Q is thermal energy; P is pressure; Y is volume.

The energy change of the selected subsystem is due to the work made by external forces. Therefore, the complete energy change of a subsystem corresponds to dE.

Through equations with the common variable of 'energy', it is possible to link classical mechanics with thermodynamics.

NOTE: *In thermodynamics, a mechanical equilibrium is defined as a uniform pressure (for a fluid). In classical mechanics, equilibrium is defined by: sum of external forces and external torques equals zero.*

4.4 Mechanical systems

A **mechanical system** is defined by its kinematics, which is described by links and coordinates. Links make up the physical composition of a mechanical system. Coordinates are used to express the time-evolution of a continuous state that results in motion. A mechanical system is defined as a collection of bodies (or links and other material components) in which some or all of the bodies can move relative to one another.

A mechanical power system produces, directs, and manages mechanical power to accomplish a task involving forces, energy transfer, and the movement of physical bodies. A mechanical power system uses forces (and energy) and the displacement of physical bodies to do mechanical work and effect change. The bodies may be rigid or non-rigid.

Mechanical work is the amount of energy transferred by a force, and mechanical power is the time derivative of mechanical work. In other words, mechanical power is a measure of the mechanical work done by means of energy transferring through a mechanical system over a certain period of time. Mechanical power is the rate of doing mechanical work; it is how fast mechanical energy is being or can be delivered to/through a mechanical system.

To clarify what is meant by mechanical work, its definition can be re-stated in a number of different ways:

1. Mechanical work is the component of the force that moves the object times the distance the object moves.
2. Mechanical work is defined as the product of the force exerted on a body and the distance it moves in the direction of that force.
3. Mechanical work is the force times the distance on which it acts.
4. Mechanical work is the force times the displacement in the direction of force.
5. Mechanical work is the action of a force moving through a distance.
6. Mechanical work is the product of a force and the displacement caused by the force when both are measured in the same direction.
7. Mechanical work is the scalar product between the applied force and the displacement vector of the motion.
8. If a force is allowed to act through a distance, it is doing mechanical work.
9. Similarly, if torque is allowed to act through a rotational distance, it is doing mechanical work.

In mechanics, the [mechanical] work done on an object is related to the forces acting on it. Mechanical work is a scalar value. A mechanical force has two attributes and two states. The two attributes are:

1. Direction
2. Magnitude

The two states are:

1. The mass of that which is displaced.
2. Its previous state of motion.

Mechanical work is equal to the force acting on an object times the distance the object is displaced (or moved). Note that only motion that is in the same direction as the force "counts". The formula for mechanical work is:

- Work (W) = Force (F) · distance (x)
- $W = \int F \cdot x$
- $W = F \times d \times \cos\Theta$
- Units of work: Joules (do not use N.m)

Differentiating by time gives indicates that instantaneous power is equal to the force times the object's velocity v(t):

- $P(t) = F(f) \cdot v(t)$

Power as a function of time, is the rate at which work is done:

- $P(t) = W / t$

Here, **acceleration** is measured by dividing an object's velocity by a unit time. On Earth, the "gravitational" acceleration constant is 9.8m/s^2, which is the rate at which an object's velocity changes. **Velocity** is equal to the distance that an object travels per unit time in a certain direction. It is a vector quantity, meaning it contains both speed and direction.

- Torque is a vector value.

A mechanical system consists of (at least):

1. A power source and actuators that generate forces and movement.
2. A system of mechanisms that shape the actuator input to achieve a specific application of output forces and movement.
3. A controller with sensors that compares the output to a performance goal and then directs the actuator input.

The 'mechanism' of a mechanical system is assembled from components called machine elements. These elements provide structure for the system and control its movement.

1. Linear kinematics
2. Angular kinematics
3. Linear kinetics
4. Angular kinetics
5. Mechanical equilibrium

DEFINITION: **Mechanical degrees of freedom (DOF)** *are classified as either scleronomic (i.e., time-independent) or* *rheonomic (i.e., time-dependent). The number of degrees of freedom (DOF) of a mechanical system is defined as the minimum number of generalized coordinates necessary to define the configuration of the system.*

4.4.1 Motion

There are four main types of motion:

1. Linear Motion – movement in a straight line
2. Reciprocating Motion – backwards and forwards or up and down movement (e.g., engine pistons and valves)
3. Angular/Rotary Motion – movement around in a circle
4. Oscillating Motion – movement over and back in an arc

There are five main types of force:

1. Tension (tensile force) – is when something is pulled and can result in stretching.
2. Compression (compressive force) – is when something is squeezed and can result in crushing.
3. Shear (shearing force) – is when something is cut or slides and results in sliding or shearing.
4. Torsion (torsion force) - is when something is twisted.
5. Bending (bending force) – is when something is bent and can be permanently deformed.

4.5 Mechanical devices

CLARIFICATION: *Commonly, a 'machine' takes in power and converts it to useful output; it is a tool containing one or more parts that transfers energy through mechanical power to perform an intended action/process. There are several types of machines: mechanical machines involving mechanical power, as well as mechanical work; computers and sensors are programmable (usually electronic or mechatronic) machines; and molecular machines (nano-machines) involve molecular components that produce quasi-mechanical movements (output) in response to specific stimuli (input); or some combination thereof.*

A **mechanical device** (or simply, **machine** or **mechanical machine**) is a system that applies mechanical power through the principles of classical mechanics to achieve desired forces and movement (motion). It is an assemblage of parts that transmit forces and motion, and transfer energy, in a predetermined manner. Machines are technological devices used to change the size, direction and speed of forces, and can also change the type of motion produced. Machines control the magnitude and direction of motion. All machines require a power source and transfer [mechanical] energy.

A mechanical device has two functions: transmitting definite relative motion and transmitting force. These functions require strength and rigidity to transmit the forces. Hence, a machine is a combination of rigid or resistant bodies, formed and connected so that they move with definite relative motions and transmit force from the source of power to the resistance to be overcome. In specific, a machine is a collection of resistant bodies arranged to change the magnitude, direction or point of application of a moving force(s) for a specific function (requirement and use). Motion is an essential part of a machine; without it, at least in principle, there is no machine, but only a structure. A **structure** transmits force without motion. A structure is a mechanism in which motion is precluded.

DEFINITION: *The* **configuration** *of a mechanical system is defined as the position of each of the bodies within the system at a particular instant. In general, both translation and rotation coordinates are needed to describe the position of a rigid body. Together the translation and rotation coordinates are called generalized coordinates. A configuration is a set containing the positions of all particles of the body.*

Complex machines involve a system of mechanisms that shape the input to achieve a specific application of output forces and movement. The term **mechanism** is applied to the combination of geometrical bodies which constitute a machine or part of a machine. A mechanism may therefore be defined as a combination of rigid or resistant bodies, formed and connected so that they move with definite relative motions with respect to one another. Whereas machines transfer energy to do work, mechanisms modify motion, and may or may not perform the function of transferring energy. In kinematics, a mechanism is a means of transmitting, controlling, or constraining relative movement. Machines may, and usually do, consist of mechanisms, with a source of power added. A mechanism is usually a piece of a larger process or mechanical system considered purely with respect to motion (kinematically). Sometimes an entire machine may be referred to as a mechanism, while still being a part of a larger machine. Examples are the steering mechanism in a car, or the winding mechanism of a wristwatch. Multiple co-joined mechanisms, however, are machines. The term **machinery** generally means machines and mechanisms.

DEFINITION: *A 'mechanism' is the fundamental physical or chemical processes involved in or responsible for an action, reaction, or other natural phenomenon.*

Mechanisms can be divided into planar mechanisms and spatial mechanisms, according to the relative motion of the rigid bodies. In a **planar mechanisms**, all of the relative motions of the rigid bodies are in one plane or in parallel planes. If there is any relative motion that is not in the same plane or in parallel planes, the mechanism is called the **spatial mechanism**. In other words, planar mechanisms are essentially two dimensional while spatial mechanisms are three dimensional.

> **NOTE:** *The restriction to resistant bodies sets fluid machines (hydraulic and pneumatic) into their own category of machine, except for a hydraulic press, which depends on statics.* **Statics** *is the branch of mechanics concerned with bodies at rest and forces in equilibrium.*

The mechanical inputs and outputs of a machine may be either forces or torques, and a machine may convert one into the other. A torque causes rotation, while a force causes linear motion. The work done is either torque times angle of rotation, or force times distance. The dimensions of torque are force times distance, and this should be carefully distinguished from work, which has the same dimensions.

A fundamental property of machines is that the input and output work are the same, except for frictional losses that make the output work smaller (the principle of energy conservation).

An "ideal machine" is one in which the parts are considered to be weightless, frictionless, and rigid. Whereas an "ideal machine" is an imagined construction, "real machines" are capable of being constructed. In practicality, there are no "ideal machines", but consideration of an "ideal machine" may aid in thought and analysis about machine design. However, take note that in some machines weight, friction, and/or flexibility/elasticity (lack of rigidity) play an essential role.

> **NOTE:** *A 'perpetual motion machine' is a machine that "works by itself and moves without applied effort from." Magnetism is one potential way of developing a perpetual motion machine. Although, over time, the magnetic force itself would run down, and it would have to be replenished.*

A **simple machine** is a machine from which no part can be removed without destroying it as a machine. A simple machine transforms the magnitudes of the forces and velocities at its point of action, but does not change the mechanical power, product of force, or velocity. **Complex machines** use more than one simple machine to accomplish a function. A complex machine is a system realized by many parts with different functions, linked together to complete a defined task. Complex machine can transform power and/or convert power from one type to another.

There are six simple machines used to control mechanical power (mechanical energy): the lever, the pulley, the wheel and axle, the inclined plane, the screw, and the wedge. Sometimes the wedge and screw are considered special cases of the inclined plane, so there are either four or six simple machines. There are an innumerable multitude of complex machines.

Simple machines can be combined in a limitless variety of ways to produce complex machines. Complex machines are, in general, composed of moving parts. Complex machines range from very basic (i.e., "simple") to highly complex. For instance, pliers, scissors and similar tools are basic-complex machines composed of two levers joined at their common fulcrum with force applied by a mechanical user.

> **NOTE:** *While not machines in themselves, the following elements are important parts of many machines (known as structural components): bearings; springs, lubricants, frames, fasteners, couplings; clutches; cams; springs; gears; and seals. They facilitate the transmission of power within [mechanical] machines.*

Non-basic complex machines may be understood in terms of system dynamics, wherein there are two classes of machine component: transformers and transducers. Transformers and transducers are interfaces that transmit power between subsystems in dynamic system models. They are the dynamic system elements that permit useful systems modelling, since most complex machines are a combination of interacting subsystems. A **transformer** is a machine element that links or interfaces two subsystems of the same type of power (energy). A **transducer** interfaces subsystems of dissimilar power (energy) types. In system dynamics, a transformer and transducer are similar in that they interface the power flow between two subsystems. Transducers differ from transformers, because they interface dissimilar subsystems. Note that the terms transformer and transducer have specific definitions in system dynamics, which differ from, but are based on, their common engineering usage.

> **CLARIFICATION:** *In mechanical engineering, the term 'transducer' most commonly refers to a sensor, such as a load cell, which emits an electrical signal in response to a non-electrical input, such as a force, in the case of a load cell. Sensors typically interface two dissimilar energy systems. In system dynamics terminology, transducers interface subsystems of different types of power (energy). There is a significant difference between a sensor and a system dynamics transducer. Sensors produce signals which are information, not power. An ideal sensor signal is time-varying voltage or current, not power, which is the product of current and voltage. Although the use of term transducer as the classification of machine elements that interface dissimilar types of energetic subsystems is logical, it is also unfortunate because it causes misunderstanding between engineers. It is best to use the term transducer only outside of system dynamics to mean sensor to avoid misunderstanding. It is also important to clarify that in most engineering communication a 'transformer' is an 'electrical transformer'. It*

is only in a systems dynamics context that the 'transformer' has a broader meaning.

Examples of transducers include:

1. DC motors interface electrical and rotational mechanical subsystems.
2. Hydraulic pistons interface fluid and translational mechanical subsystems.
3. Racks and pinions interface translational and rotational mechanical subsystems
4. Pumps interface fluid and rotational or translational mechanical subsystems

In the examples above, although the subsystems are linked by a transformer, they are also separated by it. A transformer is an interface between the subsystems that transmits power from one subsystem to another. In a transformer, the power is "transformed" twice during the energy transfer, from its original form to an intermediate form, and then, back to its original form. It is this double "conversion" of power from the power type of the subsystem to a different form and then back again, which separates the two subsystems from each other. The double "conversion" of power also permits a transformer to change the relative magnitudes of the power variables, including the magnitude of the power flow.

Actual transformers always dissipate some energy as heat due to friction, electrical resistance, or magnetic hysteresis, and retain some amount of energy. In an "ideal" transformer, all of the power which leaves one subsystem is transmitted to the other subsystem without energy storage or loss. Hence, "ideal" transformers neither dissipate nor retain energy. All the energy which flows in must flow out. Real machines with real materials have less than "ideal" links and electrical conductors. Consequently, the energetic model of machine elements that function as transformers may need energy storage and dissipation elements, in addition to the element that represents power (energy) "transformation".

The subsystems linked by transformers are comprised of energetic elements that handle the same type of power. Examples of mechanical transformers include:

- Levers and linkages, which interface translational mechanical subsystems.
- Gear sets, which interface rotational mechanical subsystems.
- Belt drives, which interface rotational mechanical subsystems.
- Double-ended pistons, which interface fluid subsystems.

Additionally, machines equipped with moving parts can be classified three ways by the type of task they perform:

1. Machines that produce mechanical power from other "forms" of energy. A combustion/electric engine is an example of a machine that produces mechanical power. If their purpose is simply to make placements or generate forces/torques, they are called actuators.
2. Machines that absorb mechanical power to accomplish a specific task (machine tools, transportation, agricultural machinery, textile machinery, machine packaging, etc.). For example, a windmill absorbs mechanical power from fluid passing through it, and a generator within the windmill machine converts mechanical power from the windmill to electrical power.
3. Mechanical transmissions: these machines transmit mechanical power by appropriately changing values of torques and speed. Mechanical transmissions are generally made up of mechanisms that have been studied (mainly from the point of kinematic view) to connect motors and users.

For example, an engine/motor is a machine in which the input is not in the form of mechanical power (i.e., it is a transducer), but it converts the power into mechanical power as forces and torques (i.e., produces mechanical power). Here, the input could be electrical (motor), or provided by a heat engine. It could also include machines worked by animate power (considered as part of the machine and not as users of it). A 'prime mover' is sometimes considered an "engine" whose power is derived from some non-mechanical source, such as a heat engine. A prime mover is capable of motion, or being moved, without connection to any other system. Windmills, water wheels and turbines are considered to be prime movers, as clearly are animate (humans and other animals).

> **NOTE:** *A* **dynamometer***, or "***dyno***" for short, is a device for measuring force, moment of force (torque), or power. For example, the power produced by an engine, motor or other rotating prime mover can be calculated by simultaneously measuring torque and rotational speed (rpm).*

Every machine has an input and an output, and the output is a modification of the input, not a simple replication of it. A machine is a processor or transformer (transducer or transceiver) in some sense. The motion of the output is fully constrained by the motion of the input, and by its kinematic connection(s). The force at the input is called the effort, and the force at the output, a load.

> **NOTE:** *Weight, as the expression of the force of "gravity" is a way of transmitting force in some machines, but it would not be considered part of a machine itself. Weight is also a common load on a machine which is not located in outer space.*

Load, in mechanics, is the external mechanical resistance against which a machine acts. For example, engine load is the power that the outside world takes away from the engine. An engine connected to nothing can have essentially no load, regardless of throttle opening or rotations per minute (RPM). If an engine is connected to a dyno or a machine, then the engine can be loaded. Therein, an engine that produces more power can accommodate more load. If the output power of the engine is less than the external load, the engine will decelerate. If the output power of the engine is greater than the external load, the engine will accelerate. Note here that the term 'load' can be confused between disciplines, because in electricity, a load is a measure of power. But in mechanics, a load is a measure of a force or a torque. In other words, power only describes load in an electrical context. In a mechanical context, it is always a force or a torque. Wherein, load is often be expressed as a curve of force versus speed.

The **mechanical advantage**, which we shall call simply the advantage, is the ratio of the load to the effort. The velocity ratio is the ratio of the movement of the load to the movement of the effort, in linear displacement or rotation. Alternatively, the **velocity ratio (or speed ration)** is the ratio of the movement of the load to the movement of the effort, in linear displacement or rotation.

> **NOTE:** *In an ideal machine the product of the mechanical advantage and the velocity ratio is* **unity***. There is a trade-off between force and speed. In a real machine the product is less than unity. As a consequence, an ideal machine in equilibrium (when the effort and the load balance) can be moved by the least impetus, as well in one direction as in the other, so the machine is reversible. A real machine, however, requires a certain effort to move it in either direction; it is irreversible, and there is an unavoidable loss of energy whenever it moves.*

4.6 Mechanical power generation

Mechanical power can be generated through the construction of a technological device that transfers/converts/transduces from fluid power (energy), electrical power (energy), thermal power (energy), magnetic power (energy), electromagnetic power (energy), or through simple inertia or gravity.

1. Fluid > linear/rotational mechanical (e.g., turbines, windmills, hydro-dam, hydraulics engine - pressurized fluid)
2. Electrical > mechanical (e.g., rotational electric motors, electrical engine)
3. Thermal > mechanical (e.g., internal combustion engines convert thermal power into rotational mechanical power - thermal engine/heat engine)
4. Chemical > mechanical (e.g., chemical motor)
5. Magnetic > mechanical (e.g., magnet)
6. Electromagnetic > mechanical (e.g., electromagnet)
7. Inertia/gravity > mechanical (e.g., sail, airfoil)

There are devices, which are not always classified as machines, but nevertheless generate mechanical power. These devices depend entirely on inertial forces and are often composed of simple machines. It is because of their reliance entirely on inertial forces that they are sometimes excluded from the definition of a machine. Such devices include but are not limited to: the pendulum; the whole family of fluid turbines; sails; and airfoils.

4.6.1 Motors and engines

Electrical > mechanical
Chemical > mechanical
Thermal > mechanical

Motors and engines convert various types of power (chemical, electrical, hydraulic, pneumatic, etc.) into mechanical power; possibly linear, but typically torque on a rotating axis. Although the terms motor and engine are often used interchangeably, they are distinguishable: engines run on thermal combustion; motors run on electricity or chemical power; and turbines run on fluid flow. There are several notable distinctions between motors and engines that may be made here:

1. A 'motor' converts electrical or chemical power into mechanical power, while an engine converts various other (non-electrical and non-chemical) forms of power to mechanical power.
2. An 'engine' is a mechanical device that uses a fuel source to create an output.
3. The word "engine" is generally used to refer to a reciprocating engine (steam or internal combustion), while "motor" is generally used to refer to a rotating device such as an electric motor.
4. An engine is made up of pistons and cylinders, while a motor is made up of rotors and stators.
5. A heat engine uses heating to generate mechanical power.

To add context, it may be useful to look at the etymological origins of the word engine and motor. The word "engine" comes from the Latin word "ingenium". An engine is a device or system (electrical, mechanical, chemical, or even social, human, or political) which effects a result. In classical mechanics, engines are basically the devices which transfer/convert energy to bring about mechanical effects. Originally, "motor" was another word for "mover" (i.e., a thing which moves the rest of the device). "Motor" did not originate from "electric motor". Historically, motors were powered by wound springs. Faraday put the word "electric" in front of "motor" to distinguish it from other motors of that time. The present-day motor, called the electric motor, is a device that transfers electrical energy through to mechanical energy. The electric motor can be broadly categorized into two classes; the AC motor and the DC motor. One could also think of engines and motors in this way: An engine is any useful man-made contrivance that takes in power and possibly raw material, and converts those into a useful mechanical output. A motor is a sub-class of engines, one that produces motive power as its primary output.

> **NOTE:** *Electrical motors and combustion engines are best suited for producing angular motion.*

Thermo-mechanical power generation systems (i.e., heat engines) use a source of thermal energy (heat source) to produce mechanical power. Thermal energy sources include: fossil, biomass, and nuclear fuels; fusion; solar; combustion; and geothermal. Power generation systems that require heat as a primary input are subject to the Carnot efficiency limitations. Hence, heat engines distinguish themselves from other types of engines by the fact that their efficiency is fundamentally limited by Carnot's theorem. Although this efficiency limitation can be a drawback, an advantage of heat engines is that most forms of energy can be easily converted to heat by processes like exothermic reactions (such as combustion), absorption of light or energetic particles, friction, dissipation and resistance. Since the heat source that supplies thermal energy to the engine can thus be powered by virtually any kind of energy, heat engines are very versatile and have a wide range of applicability.

A heat engine is a system that transfers heat to mechanical power, which can then be used to do mechanical work. It does this by bringing a 'working substance' from a higher state temperature to a lower state temperature. A heating source generates thermal power that brings the working substance to the high temperature state. The working substance generates work in the "working body" of the engine, while transferring heat to the colder "sink" until it reaches a lower temperature state. During this process some of the thermal energy is converted into work by exploiting the properties of the working substance. The working substance can be any system with a non-zero heat capacity, but it usually is a gas or liquid. During this process, a lot of heat is lost to the surroundings (i.e. it cannot be used).

> **NOTE:** *Motors and engines have to be actively held (with feedback controls) or locked in position.*

4.6.2 Turbines

Fluid > *mechanical*
Thermal > *mechanical*

A turbine (from the Latin turbo, a vortex, related to the Greek τύρβη, tyrbē, meaning "turbulence"), is a rotary mechanical device (a machine) that transfers energy from the flow of a fluid into mechanical power. A turbine is a spinning wheel that gets its energy from a gas or liquid (i.e., a fluid) moving through or past it. A turbine consists of a shaft connected to a set of blades. As the energy supply source/force moves past, and interacts with the turbine, it produces a torque through the shaft of the turbine. A turbine is a machine (energy transfer/"conversion" device) for generating rotary mechanical power from the energy of a stream of fluid (such as water, steam, or hot gas). Hence, turbines transfer energy from various types of carriers of energy into mechanical power. Essentially, a turbine transfers kinetic energy, and potential energy if the fluid is moving due to a potential difference (e.g., falling water or wind moving from high to low pressure), from the flow of the fluid into mechanical power. This mechanical power can be used for mechanical tasks, or a generator can be added to the system to convert this mechanical power into electrical power (electricity). In the case of electricity generation, turbines provide rotary (angular) mechanical power for the electric generator.

A turbine produces rotational (angular) mechanical power (mechanical energy) that may be used to generate electric power (electric potential current) via electromagnetic induction. Turbines are machines used to harness energy from fluid under pressure, and convert it into mechanical work. A turbine's mechanical power output is significantly dependent on the mechanical design of the device (e.g., blades) and the quantity of the matter (e.g., falling water) that flows through it.

Turbines have blades that spin through contact with a moving fluid material (a.k.a., "working material"). A shaft is connected to the blades that produces a torque. In other words, turbines produce torque through the rotation of a shaft connected to blades that spin due to an outside force. The power in the shaft is sometimes called "shaft power" (mechanical energy). Shaft power can be directly converted into an electrical power through the connection of an electric generator to the shaft.

> **NOTE:** *Technically, a turbine is a hydraulic mechanism because it uses the force of a liquid under pressure to work/operate.*

The mechanical power produced by a turbine can be applied directly to do mechanical work (e.g., pumping water), or it can be input into a generator to produce electrical current. Turbines have blades that spin through contact with a moving material (a.k.a., "working material"). A shaft is connected to the blades that produces a torque. In other words, turbines produce torque through the rotation of a shaft connected to blades that spin due to an outside force. The power in the shaft is sometimes called "shaft power" (mechanical energy). Shaft power can be directly converted into an electrical power through the connection of an electric generator to the shaft.

> **NOTE:** *Turbine-generator systems that produce electrical power are generally just called 'turbines'. The terminology here can be*

confusing, because technically, the turbine itself produces mechanical power, and it must be connected to an electric generator to produce (AC) electric power, but often, the turbine and generator combination are referred to as the turbine, instead of turbine-generator to clarify that the output of the system is electrical power and not mechanical power.

A turbine is a turbomachine (machines that transfer energy between a rotor and a fluid) with at least one moving part called a rotor assembly, which is a shaft or drum with blades attached. Here, the mechanical work done by the shaft is called 'shaft work'. Moving fluid acts on the blades so that they move and impart rotational power (energy) to the rotor. Early historical examples of turbines include windmills and waterwheels. In an electrical turbine, the rotor is connected to the main shaft of an electrical generator, which spins within the generator to create electricity. When a turbine is connected via a shaft to a generator (which in some cases, can be a motor that is run "backwards"), such an arrangement is called a turbo-generator.

NOTE: *Whereas windmills only do mechanical work, wind turbines generate electricity through mechanical work.*

Here, a 'working fluid' contains potential energy (pressure head) and kinetic energy (velocity head). The fluid may be compressible or incompressible. Several physical principles are employed by turbines to transfer this energy and generate power.

4.6.3 Turbine design categories

Turbines transfer the kinetic energy of fluids to kinetic energy of solids through the principle of impulse or reaction, or a mixture of the two. Hence, there are two basic types of turbine, each relying on different mechanical principles to transfer energy in a working fluid into mechanical power. While there are only two basic types of turbines (impulse and reaction), there are many variations. The basic and main difference between impulse and reaction turbine is that there is pressure change in the fluid as it passes through runner of reaction turbine while in impulse turbine there is no pressure change in the runner.

1. **Impulse turbines** - As the name suggests, an impulse turbine operates because of impulses. An impulse is a force for a very short duration. The blades of impulse turbines are impacted by the fluid, causing them to rotate in a certain direction at a considerable speed. The kinetic energy (and potential energy) of the fluid gets transferred into the rotational kinetic energy of the turbine. The resulting impulse spins the turbine and leaves the fluid flow with diminished kinetic energy (due to the transfer of some of the energy into the turbine). Essentially, an impulse turbine is a horizontal or vertical wheel that uses the kinetic energy of a fluid striking its buckets or blades to cause rotation. Newton's second law describes the transfer of energy for impulse turbines. There is no pressure change of the fluid or gas in the turbine blades (the moving blades), as in the case of a steam or gas turbine, all the pressure drop takes place in the stationary blades (the nozzles). Before reaching the turbine, the fluid's pressure head is changed to velocity head by accelerating the fluid with a nozzle. Pelton wheels and de Laval turbines use this process exclusively. Impulse turbines do not require a pressure casement around the rotor since the fluid jet is created by the nozzle prior to reaching the blades on the rotor. After turning the blades or buckets, the fluid flows out.
    - Pelton turbine (hydro turbine) - High pressure heads give rise to very fast water jets impinging in the blades resulting in very high rotational speeds of the turbine. The split bucket pairs divide the water flow ensuring balanced axial forces on the turbine runner. Pelton wheels are ideal for low power installations with outputs of 10kW or less but they have also been used in installations with power outputs of up to 200 MW.
2. **Reaction turbines** - A reaction turbine rotates due to the reaction of the fluid, either leaving or entering the turbine. In reaction turbines, the main working principle is the Newton's Third law of Motion. Reaction turbines develop torque by reacting to the gas or fluid's pressure or mass. The pressure of the gas or fluid changes as it passes through the turbine rotor blades. A pressure casement is needed to contain the working fluid as it acts on the turbine stage(s) or the turbine must be fully immersed in the fluid flow (such as with wind turbines). The casing contains and directs the working fluid and, for water turbines, maintains the suction imparted by the draft tube. Francis turbines and most steam turbines use this concept. For compressible working fluids, multiple turbine stages are usually used to harness the expanding gas efficiently. Newton's third law describes the transfer of energy for reaction turbines. Essentially, a reaction turbine is a horizontal or vertical wheel that operates with the wheel completely submerged, a feature which reduces turbulence. In theory, the reaction turbine works like a rotating lawn sprinkler where water at a central point is under pressure and escapes from the ends of the blades, causing rotation. Reaction turbines are the type most widely used.

- Francis turbine (hydro turbine) - Water flow enters in a radial direction towards the axis and exits in the direction of the axis. Large scale turbines used in dams are capable of delivering over 500 MW of power from a head of water of around 100 metres.
- Propeller and Kaplan turbines (hydro turbine) - Designed to work fully submerged, it is similar in form to a ship's propeller and is the most suitable design for low head water sources with a high flow rate such as those in slow running rivers. Designs are optimised for a particular flow rate and efficiencies drop of rapidly if the flow rate falls below the design rating. The Kaplan version has variable pitch vanes to enable it to work efficiently over a range of flow rates.

The blades (foils) of a turbine are designed for the given turbine's application and its intended working fluid. The term "foil" is used to describe the shape of the blade's cross-section at a given point, with no distinction for the type of fluid, thus referring to either an airfoil (air is the fluid) or 'hydrofoil' (water is the fluid). Also, blades may be designed for uni-directional or bi-directional functioning.

1. Uni-directional - The blades only work/function [to produce energy] in one direction.
2. Bi-directional - The blades work/function [to produce energy] in both directions.

Generally, turbines depend on the impulse of the working fluid on the turbine blades or the reaction between the working fluid and the turbine's blades to turn the turbine shaft. Several different families of turbines have been developed to optimise performance for particular fluid supply conditions.

4.6.4 Power (energy) source for turbines

NOTE: *The selection of a turbine for power generation depends largely on site conditions.*

The force to turn a **turbine** could come from a number of fluid sources. Technically, all turbines are driven in some way by the pressure/movement of a fluid; but in specific, turbines can be sub-classified by their [pressure-driven] fluid source:

1. **Fluid turbine** - fluid pressure drives the turbine. Technically, all turbines are fluid turbines, so calling a device a fluid turbine is redundant and unclear.
2. **Air [powered] turbine** - the pressure of moving air/atmosphere drives the turbine. Herein, the turbine's blades/foils are called airfoils.
3. **Wind [powered] turbine** - wind drives the turbine.
4. **Water [powered] turbine** (hydro [powered] turbine) - the pressure of moving water or another liquid drives the turbine. Herein, the turbine's blades/foils are called hydrofoils.
5. **Steam [powered] turbine** - a heat [engine] source boils water (in a boiler) results in steam, which drives the turbine. Steam turbine systems use the dynamic pressure generated by expanding steam to turn the blades of a turbine. Here, steam is produced by heat from a heat source that drives a steam turbine. Steam turbines run on the Rankine Cycle. Steam turbines rotate in the currents caused by the hot water vapour. Simply, a steam turbine system pumps liquid water into a boiler at high pressure, adds heat converting it to a super-critical fluid state and then expands it through a turbine before exhausting it to a condenser which cools it back to liquid water that feeds the pump mentioned above, creating a cycle. Generally, steam powered turbines form part of a closed water cycle in which water condenses and is then heated until it evaporates again. Steam turbines therefore do not come into contact with the fuel source that heats the water to steam. These systems generally work at temperatures between 500 and 650°C. Several steam turbines are often arranged in a row so that – configured for high, medium and low pressure – they are able to optimally convert the respective steam pressure into rotational movement.
    - Sources of heat for the production of steam include, but are not limited to: fuel combustion; solar thermal energy; waste heat; geothermal; and nuclear.
6. **Nuclear [powered] steam turbine** - boiling water (or another working fluid) from a nuclear reaction produces steam that drives the turbine.
7. **Geothermal [powered] steam turbine** - boiling water (or another working fluid) from geothermal heat transfer drives the turbine.
8. **Gas [powered] turbine** - hot gases and their expansion/combustion drives the turbine. Gas turbine systems use the dynamic (compression and then expansion) pressure from flowing gases (air and combustion products) to directly operate the turbine. Herein, expansion refers to a decrease in pressure, and increase in volume of steam or gas, which converts its pressure energy into kinetic energy (or mechanical work). Gas turbine plants run on the Brayton Cycle. Gas turbines rotate directly in the hot combustion gases. Hence, these turbine system do come into contact with the fuel and/or combustion products of the fuel. These systems reach temperatures up to 1500°C, and the gases are much hotter than those in steam turbines. For this reason the blades are generally

cooled with air that flows out of small openings and creates a "protective film" between the exhaust gases and the blades. Without cooling, the blade material would quickly wear out.
9. **Combined cycle turbines** - utilize the hot exhaust gas from the gas turbine.
10. **Combustion [powered] turbine** - the combustion of gas leading to its heated expansion drives the turbine.
11. **Osmotic [powered] turbine** - osmotic pressure drives the turbine.

NOTE: *A* **heat engine** *is a power production system that converts heat (thermal energy) to mechanical power (mechanical energy), which can then be used to do mechanical work (e.g., powering a turbine).*

All thermal sources of mechanical power produce "waste heat" as a by-product of the useful mechanical power produced. The percentage of heat transferred into useful mechanical power (mechanical energy) is know as 'conversion efficiency'. It is not thermodynamically possible for all of the heat (thermal energy) to be transferred into mechanical power, according to the second law of thermodynamics; therefore, there is always heat lost to the environment. If this loss is employed as useful heat, for a separate heating service, the power generating system is referred to as a **cogeneration power plant** or **CHP (combined heat-and-power) plant**. By-product heat can be used for multiple purposes, including: heating the exterior and interior of architecture, and for the desalination of water. Dedicated heat plants called **heat-only boiler stations** do not produce electrical power, but instead generate thermal energy in the form of hot water for use in heating applications. Heat-only boiler stations can generate heat from the same sources that generate electrical power.

4.6.5 Biomechanical power generators

In general, animate power is considered a form bio-electric-mechanical power. Animals produce "animate" power through at least bio-electric means, which drive their appendages to produce mechanical power. Animate power may produce a pushing, pulling, or torquing power. This power can then be harnessed to do useful work. For instance, a horse can be used to for its "horsepower" to pull a cart. A human can operate a hand crank. Or, an animal can run in a wheel.

4.7 Mechanical power transmission

Mechanical Power transmission is the movement of energy from its mechanical place of generation to a location where it is applied to performing useful work. Mechanical power transmission refers to the transmission of motion and power from generation (a driver or source) to use (a load or output). Therein, mechanical power may be transmitted:

1. Within a mechanical device.
2. From a mechanical device to/through its end-point of use (load or output).

Typically, a transmission element will have an input side and an output side, and the motion from input to output is related, assuming no losses, in one of two ways.

1. Geometric constraint: For a transmission element, there will be a relationship between the geometry of the motion at the input to the geometry of the motion at the output.
2. Energy conservation constraint: For an ideal transmission element, the power input to the element will be equal to the power output from the element.

The geometric constraint must hold if the transmission element is operating properly. For example, the speed of the motion at the interface between two meshing gears must be the same or the teeth will be sheared off the gears. The energy conservation constraint holds if the transmission element does not have significant energy dissipation. Typically, this is approximately true for a good transmission element, because it is designed to transmit as much of the input power to the output as possible. Of course, neither constraint is strictly true in reality, but, if deviations are small, these can be incorporated into other ideal elements in the system model.

A linkage can be used to change the direction of a force or to make two or more things move at the same time. A **mechanical linkage (**or **link)** is an assembly of bodies connected to manage forces and movement. A link is a mechanical part that transmits an axial force of compression or tension, and is connected by pins or sliders at its ends. A linkage can be used to change the direction of a force or to make two or more things move at the same time. A link is not a machine by itself (it does not transform its input), but is a typical part of a mechanism, and may transmit forces between simple machines. In general, it involves a rigid body having two or more pairing elements which connect it to other bodies for the purpose of transmitting force or motion. A slotted link with a sliding block may permit a variable amount of motion to be transmitted. In every machine, at least one link either occupies a fixed position relative to the earth or carries the machine as a whole along with it during motion. In the later case, the link is the frame of the machine, and it is called the fixed link. The movement of a body, or link, is studied using geometry; hence, the link is considered to be rigid. The connections between links, which are called **joints**, and are modelled as providing ideal movement, pure rotation, or sliding. The combination of links and joints without a fixed link is not a mechanism, but a 'kinematic chain'. In other words, a linkage modelled as a network of rigid links and ideal joints is called a 'kinematic chain'.

Linkages may be constructed from open kinematic chains, closed kinematic chains, or a combination of open and closed chains. Each link in a chain is connected by a joint to one or more other links. Thus, a kinematic chain can be modelled as a graph in which the links are paths and the joints are vertices, which is called a **linkage graph**.

Mechanical linkages are usually designed to transform a given input force and movement into a desired output force and movement. The ratio of the output force to the input force is known as the **mechanical advantage** of the linkage, while the ratio of the input speed to the output speed is known as the **speed ratio (velocity ratio)**.

Mechanical power can be transmitted across distances in a variety of ways. Additionally, mechanical power can be transmitted indefinitely, given sufficient power, as well as adequate gear boxes and axles. In general, mechanical power transmission is accomplished in one of five categorical ways, called 'drives':

1. **Belt drives** - Power transmitted through the use of belts under tension between two or more sheaves or pulleys.
2. **Chain drives** - Power transmitted through a chain between two or more sprockets.
3. **Gear drives** - Power transmitted through two or more mating gears. Can be either open (exposed) or enclosed (gears inside a gear box or reducer).
4. **Hydraulic drives** - Power transmitted through fluid pressure.
5. Some combination of the above.

Below are the delineated mechanisms (non-hydraulic) by which mechanical power can be transmitted:

1. **Clutches** - A clutch is a mechanical device that connects the power source to the rest of the machine. A clutch is used in motor vehicles so that the engine can remain running while the car is at rest, start slowly without stalling, and shift gears while moving.
2. **Pulleys and belts** - belts and pulleys control mechanical energy through any of 5 different arrangements: connect and disconnect power like a clutch, change direction, reverse rotation, change speed, and change torque.
3. **Chains and sprockets** - usually used as the drive system to bring power to the driving wheel of the vehicle in bicycles and mopeds.
4. **Shafts and bearings** - used to transfer mechanical energy in many types of machines. Shafts transfer motion (mechanical power) from point to point along their axis of motion.
5. **Gears** transfer power between shafts while keeping fixed ratios between shaft speeds. Gears transfer motion (mechanical power) via their contact point's pitch diameter. In other words, the pitch diameter is the point in which both gears transfer power. The torque from the gears transfer power from one part of a machine to another.
6. **Rails** for conveying suspended objects.

Gears are not just used to transfer power, they also provide an opportunity to adjust the mechanical advantage (the ratio of output force to the input force) of a mechanism.

Below is a delineation of the common types of gears:

1. **Spur gears** transfer motion between two shafts running parallel to each other.
2. **Bevel gears** are conically shaped, and transmit power between shafts that have intersecting axes of motion.
3. **Crown gears** mesh with bevel gears and spur gears so that motion is transferred between shafts with intersecting axes of rotation.
4. **Worm gears** come in pairs (worm gears and worm wheels) combine to transfer power between perpendicular shafts that have axes of rotation offset from each other.
5. **Helical gear**s transmit power between two parallel axes of motion, or between perpendicular non-intersecting axes of motion. Helical gears resemble spur gears, only their teeth are curved in the shape of a helix.
6. **Epicyclic** or **planetary gear** sets consist of one or more planet gears moving along an outer ring gear as a central sun gear drives them. As the planet gears are driven, they typically move a planet carrier plate along with them. The overall mechanical function of a planetary gear set changes depending on the configuration used.
7. **Rack and pinion gear sets** are used to convert rotational motion to linear motion. A rack gear is a gear mounted to a straight rod, such that it moves in a linear fashion when torque is applied to it by a spur gear (known as the pinion gear).

A **transmission** is a machine in a mechanical power transmission system that provides for the controlled application of power. Often the term transmission refers simply to the gearbox that uses gears and gear trains to provide speed and torque conversions from a rotating power source to another device.

> **NOTE:** *An "ideal" transmission element transforms one type of motion/force/moment into another without a loss of power. That is, in an "ideal" transmission element, there is no loss of energy.*

The disadvantages of mechanical power transmission include, but are not limited to:

1. Lubrication problems
2. Limited speed and torque control capabilities
3. Limited transmission distance
4. Uneven force distribution
5. Physical space requirement

4.8 Mechanical power production systems

These are systems that produce mechanical power, either through another mechanical power input (e.g., wind), or through the conversion of power from another carrier (e.g., electrical power).

1. **Magnetic [field] motor systems** - The use of permanent and/or electromagnets to produce mechanical power. There are the three types of permanent magnetic motor system:
   - The imbalanced system (spin alignment system)
   - The induction expulsion system (catch and release)
   - The exchange force pulse system (spin accelerator system)
2. **Electric motor systems** - The use of electrical power and electromagnetic induction to produce mechanical power. A motor is a machine that converts electrical power (electrical energy) into mechanical power (mechanical energy). Electric motors are used to produce linear or rotary force (torque), and should be distinguished from devices such as magnetic solenoids and loudspeakers that convert electricity into motion but do not generate usable mechanical powers, which are respectively referred to as actuators and transducers. In a motor, rotational mechanical power (torque) is transferred through a rotor shaft. Energy loss during motor operation is dissipated as heat, so they sometimes have fans to cool down the motor. There exist both AC and DC electric motors.
   - The electric motor could be called an electromechanical continuous energy conversion device.
3. **Combustion engine systems** - The use of chemical power as combustion to produce mechanical power.
4. **Turbine systems** - The use of the mechanical power of a moving fluid to produce mechanical/ fluid power.
5. **Transducers** - The use of electrical power to produce mechanical power (or vice versa) as the vibration of a medium.

5 Fluid power systems

> **Fluid mechanics** *is a branch of classical mechanics (a branch of physics) that studies the mechanics (and dynamics) of fluids (liquids, gases, and plasmas) and the forces on them (i.e., the behavior of fluids). Fluids at rest are known as hydrostatics, and fluids in motion are known as fluid dynamics. Dynamics divides into two branches depending on the consideration of the viscosity to describe the flow (inviscid flow is where the influence of viscosity is neglected, and viscous flow considers viscosity as a dominant parameter that influences flow.*

Fluid [mechanical] power is the use of fluids under pressure to generate, control, and transmit power. Take note that there are two principal types of mechanical system, solid and fluid. A **fluid** (and liquid) is a substance that deforms continuously when a shear stress is applied. Both liquids and gases are fluids. Fluid Power is produced by outside energy sources, such as a motor. The fluids transmit the energy, and are not the source of fluid power. A fluid is a material that can flow, has no definite shape of its own, and conforms to the shape of its container. Fluids spontaneously move from regions of high pressure to regions of low pressure. Fluid power transfers energy through the variables of pressure and flow. Flow is necessary for the development of pressure, which is a function of resistance to fluid flow in the system. Liquids in motion have characteristics different from liquids at rest. Frictional resistances within a fluid (viscosity) and inertia contribute to these differences.

In a fluid power system, the pressure is typically transferred to some type of actuator used to perform work. Actuators can be rotary, linear or a combination of the two. Linear actuators are often referred to as cylinders or rams, while rotary actuators are called motors.

Fluid power is subdivided into hydraulics (using a liquid such as mineral oil or water), and pneumatics (using a gas such as air or other gases). At a very basic level, hydraulics is the liquid version of pneumatics, or said in the opposite way, pneumatics is the gases version of hydraulics.

1. **Hydraulics** is the study of liquids at rest and in motion -- the study of the motion of liquids in relation to disciplines such as fluid mechanics and fluids dynamics. The science and engineering of forces and movement transmitted by means of liquids. The word “hydraulics” originates from the Greek word δραυλικός (hydraulikos), which in turn originates from ὗδωρ (hydor, Greek for water) and ανλός (aulos, meaning pipe).
2. **Pneumatics** is the study of gases and their behavior under pressure. The science and engineering of forces and movement transmitted by means of gases or pressurized air. Pneumatic

power is a measure of work produced using pressurized gases/air. The principles of pneumatics are the same as those for hydraulic, but pneumatics transmits power using a gas instead of a liquid.

Fluid systems (hydraulics and pneumatics) are best suited for producing linear motion.

Hydraulics is a topic in applied science and engineering dealing with the mechanical properties of liquids and fluids. Fluid mechanics provides the theoretical foundation for hydraulics, which focuses on the engineering uses of fluid properties. In fluid power systems, hydraulics are used for the generation, control, and transmission of power by the use of pressurized liquids. Hydraulic topics range through some part of science and most of engineering modules, and cover concepts such as pipe flow, dam design, fluidics and fluid control circuitry, pumps, turbines, hydropower, computational fluid dynamics, flow measurement, river channel behavior and erosion.

Hydraulic power is a measure of the work produced by putting liquids under pressure, and their consequential flow. There are three types of energy available in modern hydraulics (of the normal hydrostatic type):

1. **Potential energy & pressure energy** - The static energy of a standing, but pressurized liquid that is ready to do work (e.g., oil in a loaded accumulator).
2. **Kinetic energy** - The energy of the moving liquid, which varies with the velocity (speed) of the liquid.
3. **Heat transfer energy** - Friction or resistance to flow (an energy loss in terms of output). Example: friction between moving oil and the confines of lines or passages produces heat energy.

Hydraulics and hydro-mechanics engineering science of liquid pressure and flow. There are two branches of hydraulics/hydro-mechanics:

1. **Hydrodynamics** - The engineering science of the energy of liquid pressure and flow - dynamic effect through mass times acceleration. Force effect through pressure area.
2. **Hydrostatics** - The engineering science of the energy of liquids at rest - dynamic effect through pressure times area. Force effect through mass acceleration
3. **Free surface hydraulics** is the branch of hydraulics dealing with free surface flow, such as occurring in rivers, canals, lakes,estuaries and seas. Its sub-field open channel flow studies the flow in open channels. It is part of the field of hydrology.

In general, fluid power systems involve a pump driven by a prime mover (such as an electric motor or internal combustion engine) that transfers mechanical power (energy) through to fluid power (energy).

All fluid systems have two things in common. First, each system contains a fluid – either a liquid or a gas – that moves through a system of connecting pipes and devices. Second, a pressure difference in the system creates a net force, which causes fluids to move or perform some special function – like pushing a piston or opening or closing a valve. In this sense, pressure is a prime mover in fluid systems.

> **NOTE:** *Fluid pipe networks and electrical wire networks are analogous. For instance, an adjustable water tap for a home water supply is just like a variable electrical resistor.*

Fluid systems are either of an 'open' or 'closed' type:

1. Open systems move fluids into and out of the system, without retaining or recirculating fluids.
   - Work occurs in a fluid system when a fixed volume of fluid V moves through a pressure difference (ΔP). If the pressure increases, ΔP is positive and W is negative. A negative value for work means that work is done on the fluid (e.g., a pump for a city's water supply, irrigation system, fire truck water system). If the pressure decreases, ΔP is negative and W is positive. A positive value for work means the fluid does the work (e.g., a hydroelectric dam). In a hydroelectric dam, water flows from a high-pressure region behind the dam to a low-pressure region, turns a turbine, and does work.
   - The formula is: $W = \Delta P \times V$
2. Closed systems retain and recirculate fluids.
   - Work occurs in a fluid system when fluid pressure p causes a given volume (ΔV) of liquids or gases to move. Here, ΔV can be positive or negative, but P is always positive. If the fluid volume increases, ΔV is positive and W is positive. Positive work means the fluid does work, as when a gas expands in a cylinder, lifting a load (other examples include: a hydraulic lift; a hydraulic break system; the body's circulatory blood system). If the fluid volume decreases, ΔV is negative and W is negative. A negative value for work means that work is done on the fluid (e.g., a weight or force applied to the piston compresses the gas in a cylinder, a scuba tank compressor).
   - The formula is: $W = P \times \Delta V$

Hydraulic and pneumatic power is maintained through a combination of fluid flow and pressure. When discussing fluid power, pressure is the basis for producing any kind of work. Work cannot be achieved without pressure. **Pressure** is defined as the measure of force acting perpendicular to a unit area. Force is anything that tends to produce or modify (push or pull) motion. Pressure is applied in all directions regardless of shape or size. Pressure can act both outward and inward, depending on the circumstances. Additionally, pressure will always act perpendicular to the surface of the body upon which it is acting.

- Force (F) = pressure (P) · area (A)
- Pressure (P) = Force (F) / area (A)
- Fluid pressure (P) = force (F) / unit area (A)
- Fluid flow rate (Q) = volume (V) / unit time (A)
- Fluid power = pressure (P) x flow rate (Q)

The elements of an electrical system are analogous to a fluid system:

- Pressure = voltage
- Volume = capacitance
- Flow rate = current
- Flow restrictions = resistance
- However, air is unlike electricity in that air is compressible. Hence, the elements of a fluid system have more non-linearities than those of electrical systems.

Pressure and flow are essential design considerations for a fluid power system. **Pressure** refers to matter pushing against matter. For instance, an object pushing against another object.

- Absolute (psia) - true matter based pressure.
  - 0 psia - no matter present to press against object(s)
- Gage (psig) - relative to atmosphere.
  - 0 psig - pressure in equilibrium with atmosphere.

Flow is a loose term that generally has three distinct meanings:

1. **Volumetric flow** is used to measure volume of fluid passing a point per unit of time. Where the fluid is a compressible gas, then temperature and pressure must be specified or flow normalised to some standard temperature and pressure.
2. **Mass flow** measures the mass of fluid passing the point in unit time.
3. **Velocity of flow** measures linear speed past the point of measurement. Flow velocity is of prime importance in the design of hydraulic and pneumatic systems.

The most important physical properties of fluids are:

1. **Density** - can be considered constant.
2. **Viscosity** - varies greatly with temperature and less greatly with pressure.

5.1 Comparison between pneumatic systems and hydraulic systems

The fluid generally found in pneumatic systems is air; in hydraulic systems it is oil (or water). And, it is primarily the different properties of the fluids involved that characterize the differences between the two systems:

1. Air and gases are compressible, whereas oil is incompressible (except at high pressure).
2. Air lacks lubricating property and always contains water vapor. Oil functions as a hydraulic fluid as well as lubricator.
3. The normal operating pressure of pneumatic systems is very much lower than that of hydraulic systems.
4. Output powers of pneumatic systems are considerably less than those of hydraulic systems.
5. Accuracy of pneumatic actuators is poor at low velocities, whereas accuracy of hydraulic actuators may be made satisfactory at all velocities.
6. In pneumatic systems, external leakage is permissible to a certain extent, but in thermal leakage must be avoided because the effective pressure difference is rather small. In hydraulic systems internal leakage is permissible to a certain extent, but external leakage must be avoided.
7. No return pipes are required in pneumatic systems when air is used, whereas they are always needed in hydraulic systems
8. Normal operating temperature for pneumatic systems is 5° to 60°C. The pneumatic system, however, can be operated in the 0° to 200°C range. Pneumatic systems are insensitive to temperature changes, in contrast to hydraulic systems, in which fluid friction due to viscosity depends greatly on temperature. normal operating temperature for hydraulic systems is 20° to 70°C.
9. Pneumatic systems are fire- and explosion-proof, whereas hydraulic systems are not, unless non-flammable liquid is used.

5.2 Hydraulic power generation (sources)

Most fluid power systems involve a pump driven by a prime mover (such as an electric motor or internal combustion engine) that converts mechanical power into fluid-hydraulic power. Hydraulic power can also be generated through inertia (e.g.,a stream of water or weight of an object).

When a hydraulic pump operates, it performs two functions. First, its mechanical action creates a vacuum at the pump inlet which allows atmospheric pressure to force liquid from the reservoir into the inlet line to the pump. Second, its mechanical action delivers this liquid to the pump outlet and forces it into the hydraulic system.

> **NOTE:** *A pump produces liquid movement or flow: it does not generate pressure. It produces the flow necessary for the development of pressure which is a function of resistance to fluid flow in the system.*

Theoretical hydraulic power is calculated as :

- Hydraulic Power (Watts) = Pressure (Pa) x Flow (m3/s)
- In order to calculate hydraulic power in the units normally use in hydraulics, this formula is modified to:
- Hydraulic Power (kW) = Pressure (Bar) x Flow (l/min) / 600

5.2.1 Hydraulic mechanical and electrical power generation

Hydraulic power systems can be designed to produce mechanical power through the application of a hydraulic motor, and electrical power by connecting the hydraulic motor to a generator, thus creating a hydraulic generator. A 'hydraulic generator' converts the hydraulic power of a "working machine" into electrical power. A "working machine" can be equipped with a hydraulic generator, to generate power for itself. A hydraulic generator uses the power of a working machine's hydraulics to turn a generator and produce electrical power.

The hydraulic piston of a working machine can be connected to a set of wheels and gears that transform the translation into rotation, and speed it up. This is called an hydraulic motor. A hydraulic motor is then connected to the generator.

There are some hydroelectric machines that use hydraulics to generate electrical power. For example, the Pelamis wave energy converter used hydraulics to generate electricity from waves.

5.3 Hydraulic power transmission and distribution

Most hydraulic power is transmitted via a cylinder, or via pipe (a.k.a., tube or hose) network. In hydraulic transmission of power, a pump is used to raise the pressure of a liquid (most commonly, oil) and energy in the liquid is transmitted through pipes and hoses to perform useful work. Pipes are suitable for power transmission over intermediate distances; they can be employed over greater distances than mechanical types of power transmission, but not as great as electrical power transmission systems. In order to control the transmission of hydraulic power from the pump to the actuators a range hoses, tubes, and possibly control valves are used. The speed of a motor or cylinder and the torque or force that can be generated is infinitely controllable using directional, flow and pressure control valves.

Hydraulics pumps, motors, and cylinders are "power dense" in that the amount of power they can absorb and transmit provides designers of machines the flexibility to locate the pumps and actuators in the most advantageous position.

Hydraulic systems can be designed to give fast operative power and move heavy loads. They can easily generate linear motion using liner actuators (also called cylinders). Speed control is simple, and precise motion of the actuator is possible.

In general, hydraulic systems use a incompressible fluid, such as oil or water, to transmit forces from one location to another within the fluid.

Hydraulic fluid(s), also called hydraulic liquid(s), are the medium by which power is transferred in hydraulic machinery. Some hydraulic systems work most efficiently if the hydraulic fluid used has zero compressibility.

The disadvantages of hydraulic system include fluid leakage, containments and fire hazards with flammable hydraulic fluids.

One of the key advantages of hydraulic systems is to be able to transmit large amounts of power from a remote power source (electric motor or internal combustion engine) to a compact actuator.

> **NOTE:** *Hydrostatic transmission (hydrostatic drive) - the transmission of mechanical power by pressurizing and releasing fluid through specialized pumps.*

5.4 Pneumatic power generation

Most fluid power systems involve a pump or other pressurizing device (e.g., air compressor) driven by a prime mover (such as an electric motor or internal combustion engine) that converts mechanical power into fluid-pneumatic power.

Pneumatic power may also be generated from:

- A trompe is a water-powered gas compressor, commonly used before the advent of the electric-powered compressor. A trompe is somewhat like an airlift pump working in reverse.
- A bleed air systems on an engine.

Pneumatic power system are often capable of, and work through, both pumping and vacuum action.

> **NOTE:** *Air has some basic an important properties. First, it is compressible. Second, if there is higher pressure, then there is higher friction. Third, the 'Ideal Gas Law': PV = nRT. Where, pressure is proportional to temperature (T), and pressure is inversely proportional to volume (V).*

5.5 Pneumatic power transmission and distribution

Most pneumatic power is transmitted via a cylinder or pipe (a.k.a., tube or hose) network, in the same way that hydraulic power is transmitted. In general, pneumatic systems are used for the transport of objects.

All gases are readily compressible and it is this property which differentiates them most from liquids as a power transmission medium. In pneumatic transmission of energy, a compressor is used as the power source to raise the pressure of the air to the required level quite slowly. They are suitable for power transmission over intermediate distances. Pneumatic systems use simple equipment has small transmission lines, and do not present a fire hazard. The disadvantages of pneumatic system include a high fluid compressibility and a small power to size ratio of components. Pneumatic systems are unsuitable for uniform motion. Operating pressure of pneumatics is around 6 to 8 bar. And hence are capable of generating only medium forces. The switching time of control elements is usually greater than 5 milli seconds and the speed of the control signal is 10 to 50 m/s. Table 1.5 give the comparison of all the systems.

The two primary types of pneumatic power transmission system are: pneumatic conveyors and pneumatic tubes:

1. Pneumatic conveyor - A pneumatic conveyor essentially comprising a tubular channel, the lower part of which defines a rail for guiding and propelling objects to be transported. Generally, the tubular channel is equipped at regular intervals with a means for guiding and supporting pressurised air conduits and electrical wiring, and by covers, which can be click-locked onto the guide and support means and which cover the corresponding edges of the tubular channel.
2. Pneumatic tube system - Pneumatic tube systems (also called PTT, airlift, air transport, Lamson tubes, air tubes, and pneumatic transit systems) has a compressed air pump attached that can either suck air from the tube or blow air into it according to which way down the tube packages need to be sent. This means the compressor may be working like a vacuum cleaner so it sucks air along the tube from the sending station. As the compressor sucks on the tube, it creates a partial vacuum in front of a canister within the tube that sucks it all the way along until it reaches a receiving station, where it can be unloaded. Canisters can be sent in the opposite direction simply by setting the compressor to blow air along the tube in the opposite direction (behind a canister, pushing it along). Just as a vacuum cleaner is limited by the suction power of its electric motor, so pneumatic transport tubes are limited in what they can carry, how quickly, and how far. Most pneumatic tube systems are very simple networks linking one receiving station with a number of sending stations, or vice-versa. However, much more elaborate, computer-controlled systems are also commonplace, in which many sending stations link to many receiving stations and packages can route and transfer in all manner of complex ways; these are the sorts of systems that hospitals or office buildings use. A large pneumatic system might have up to 500 sending and receiving stations, dozens of transfer units where packages can be routed between senders and receivers in complex ways, and dozens of compressor/blower units to provide the pneumatic power. Pneumatic tube systems are a fast, simple, secure, and reliable way of transporting small objects relatively large distances across a building, or (using underground or overground pipes) between buildings (or within a city) on the same site. They can move things up, down, or sideways and, because they're pneumatic, they provide a soft, air-cushioned ride for fragile items (many systems use air-cushioned brakes or bumpers that bring arriving canisters slowly to a rest at the receiving station).

5.6 Pneumatic energy "storage"

Pneumatic power (energy) may be "stored" in the following ways:

1. Storage Tanks
2. Tubing, Fittings & Valves
3. Compressor

6 Electrical power systems

TERMINOLOGY:

1. **Electric** is used to describe things pertaining to the set of physical phenomena associated with electric charge (i.e., "electricity").
2. **Electrical** can be used nearly everywhere that electric is used when pertaining to "electricity" and the study or application of electric charge (aside from some set phrases). For instance, generally, people do not say "electric engineer" unless the engineer runs on "electricity"; instead they say "electrical engineer".
3. **Electricity** is the set of physical phenomena associated with the presence and flow of electric charge; it is not a single thing.
4. **Electronics** refers to technology that works by controlling the motion of "electrons" and electric charges in ways that go beyond electrodynamic properties like voltage and current. Electronics is a field of science and branch of engineering. Electronic technologies are powered by electrical charges (and "electrons"), and composed of electrical circuits that involve active electrical components and associated passive interconnection elements. Electronic devices make use of the transistor as a fundamental building block of all modern electronics circuitry. A modern integrated circuit may contain several billion miniaturised transistors in a region only a few centimetres square.

'Electric power' refers to the "flow rate of electrical energy" or "rate of doing electric work". Remember that power is the flow rate of energy (i.e., rate of energy transfer), or a rate of use of energy (i.e., rate of doing work). Energy is measured in Joules, and when energy flows, the flow is measured in Joules per second (or watts). The word "Watt" is just another way of saying "Joule per Second" -- it is a unit measurement of electrical power. Energy comes in Joules, while power comes in Joules per second (Watts). In other words, the SI unit of electric power is the Watt. The term wattage is used colloquially to mean "electric power in watts".

> **NOTE:** *Electrical energy can be "stored", but electric power is not something that is ever stored. Think in this way: we can store a volume of water, but it's impossible to store any "volumes per second" of water.*

Electrical power is a convenient way to transfer energy, and to manage its transfer through useful service. Unlike hydrocarbon fuels, electrical energy is a low entropy form of energy that can be converted into motion (and many other forms of energy) with high efficiency, as well as provide the ability to store and process information at high efficiency. Through the use of technology, electrical power (energy) can be converted/transformed into light, thermal energy (heat), mechanical energy (macro- and micro-scopic motion), and other carriers of energy. Additionally, the control of electrical power (energy) allows for the creation and reception of electromagnetic radiation (e.g., radio waves) for communication and wireless electric power transmission.

Electric power exists where electric current is applied to "energise" electrical technology. All electrical power depends on an energized conductor and a path to ground. It is a path between the two that creates the flow of energy as electricity, through items that use it.

> **CLARIFICATION:** *To '***energize***' means to supply voltage (the force of electrical pressure) through a circuit. To '***de-energize***' means to remove the supply of voltage from some part of the circuit. A circuit can be de-energized through the addition or movement of a voltage isolator, circuit breaker, on/off switch, or the removal of a fuse or transmissive link; whereby, either no electrical current, or no electrical current at the requisite voltage, can flow to or from the transmission system through the de-energized connection point to the load or end-use. Conversely, to energize means to change the state of the circuit through a change in the state of a connection point so as to enable electrical current to flow to or from the transmission system to the load or end-user.*

Electrical power (energy) is "supplied" by the combination of electric current and electric potential delivered by a circuit. At the point that the electric potential energy has been transferred to another carrier of energy, it ceases to be electric potential energy. Thus, all electrical energy is potential energy before it is delivered to the end-use. Once converted from potential energy, electrical energy can always be called another "type" of energy (heat, light, motion, etc.).

Electric current flows from the state of higher potential charge to the state of lower potential charge. This flow is most often called electric power, but it may also be called "electrical energy" flow or "electricity". Therein, electric power is the product of two electric quantities, voltage and current -- there must be both voltage and current for electrical power (energy transfer) to be present:

1. **Current** is the rate at which charge is flowing. Amperage (I) is the rate at which current flows through a conductor. The single unit is the ampere. Note here that current is different than power. Current is the rate at which charge is flowing, and power is the rate at which energy is flowing.
2. **Voltage** is the difference in charge between two points. Voltage (E) is the pressure that pushes

current through a conductor. The single unit is the volt.

NOTE: *In Steinmetz' electrical theory, the magnetism (Φ) and dielectricity are the two components of electricity. It is the product of those two quantities. If it is one or the other, then it is not electricity. For instance, a charged capacitor is not electricity.*

NOTE: *Electricity does not flow in conductors. Metals used to be called "non-electrics" because they destroyed the electric field. In metallic electrical conductors (e.g., wires), surface charges accumulate along the wire, which maintain the electric fiend in the direction of the wire. Note that metallic electrical conductors maintain a surface charge distribution since any extra charge on a conductor will reside on the surface. It is the change in, or gradient of, the surface charge distribution on the wire that creates, and determines the direction of, the electric field through a wire or other resistor. In a DC circuit, the surface charge density on the wire near the negative terminal of a battery, for instance, will be more negative than the surface charge density on the wire near the positive terminal. The surface charge density, as you go around the circuit, will change only slightly along a good conducting wire - the gradient is small, and there is only a small electric field. Corners or bends in the wire will also cause surface charge accumulations that make the electrons flow around in the direction of the wire instead of flowing into a dead end. Resistors inserted into the circuit will have a more negative surface charge density on one side of the resistor as compared to the other side of the resistor. This larger gradient in surface charge distribution near the resistor causes the relatively larger electric field in the resistor (as compared to the wire). The direction of the gradients for all the aforementioned surface charge densities determine the direction of the electric fields.*

These two quantities (voltage and current) can vary with respect to time (AC electrical power), or they can be kept at constant levels (DC electrical power). Note that with AC power, both voltage and current are changing in sign (+ / -) each half cycle.

As an expression, electrical power in watts is produced by an electric current (I) consisting of a charge of (Q) coulombs every (t) seconds passing through an electric potential (voltage) difference of (V):

- P = work done per unit time = VQ / t = VI
- V is electric potential or voltage in volts
- Q is electric charge in coulombs
- t is time in seconds
- I is electric current in amperes

Hence, electrical power (P) is delivered by a combination of voltage (V) and current (I):

- power (P) = volts (V) • current (I)

In other words, the electrical power (P) delivered to a component is given by:

- P(t) = I(t) • V(t)
- P(t) is the power, measured in watts
- V(t) is the potential difference (or voltage drop) across the component, measured in volts
- I(t) is the current flowing through it, measured in amperes
- (t) refers to any point in time

There is a difference in how power is determined between direct current (DC) and alternating current (AC) circuits. In a DC system, P = VI, but in an AC system, power is a complex quantity involving the concept of a 'power factor'. There is no 'power factor' in DC circuits because there is no concept of phase angle between current and voltage. The concept of a 'power factor' only arises when voltage and current has a phase difference. It may be said that the power factor in a DC circuit is always 1, because there are no reactive components. The current and voltage are always in phase, which is another way of saying that the phase difference between the current and voltage is zero degrees (0˚), and the cosine (cos) of zero (0) is one (1).

In an AC circuit, the ratio of 'active power' to 'apparent power' is called the **power factor**. In other words, the 'power factor' in an AC circuit is the ratio of the power utilized by load to the 'total power' supplied. In general, the power in a AC circuit (single phase) is the average power (i.e., real power), which is given by:

- P^{avg} = VI cosφ = voltage x current x power factor
- φ is the phase angle between the current and the voltage.
- cos φ is the 'power factor' of the circuit.
- Power factor = Active Power/Total Power

The power factor is one when the voltage and current are in phase. If voltage and current are in phase, then power sine wave(V*I) will always be in one direction when V and I are multiplied at any time-frame, meaning power is utilized through the load. It is zero when the current leads or lags the voltage by 90 degrees. Where the waveforms are purely sinusoidal, the power factor is the cosine of the phase angle (φ) between the current and voltage sinusoid waveforms. Therein, power factors are usually stated as "leading" or "lagging" to show the sign of the phase angle of current with respect to voltage. Voltage is designated as the base to which current angle is compared, meaning that we think of current as either "leading" or "lagging" voltage. This means that power is not being utilized entirely, and that the load tends to act as a source of power. The positive area of power is the power utilized (active power) and the total area is the total power you supply.

The power flow in an AC electrical system has five components (and each is assigned a different unit of expression):

1. **Active power** or **real power** or **true power (P^{avg})**: The average value of power. Power that is actually consumed. Power that does/performs work. In a vector diagram, it is the real axis. Expressed in watts (W).
   - Real power = P = |V| | I | cos φ
2. **Reactive power (Q)**: Also known as "use-less power" or "wattless power" is the powers that continuously bounce back and forth (oscillates) between source and load. This power does not perform work. Expressed in volt-ampere-reactive (VAR). Reactive power represents energy that is first stored in the load, and then released in the form of a magnetic field (in the case of an inductor) or electrostatic field (in the case of a capacitor). In other words, reactive power is power that is stored in components, then released again back to the source through the AC cycle. Capacitors and inductors both do this, just in opposite phase. Reactive power does not do any work, so in a vector diagram it is represented as the imaginary axis of the vector diagram.
   - Reactive power = Q = |V| | I | sin φ
3. **Complex power (S)**: The vector sum (in a vector diagram) of active and reactive power. Expressed in volt-ampere (VA).
4. **Apparent power (|S|)**: The magnitude of complex power S -- the magnitude of the vector sum of active and reactive power. It is the total power in an AC circuit, both dissipated and absorbed/returned. Expressed in volt-ampere (VA). Apparent power is the product of the root-mean-square of voltage and current. The peak voltage times the peak current (or the RMS voltage times the RMS current, depending on if you're looking at peak power or average power). A power supply must be capable of outputting the full apparent power delivered to a circuit, not just the active power.
   - $|S| = \sqrt{(P^2 + Q^2)}$
5. **Phase of voltage relative to current (φ)**: the angle of difference (in degrees) between current and voltage; current lagging voltage (quadrant I vector), current leading voltage (quadrant IV vector). In other words, this is the angle used to describe the phase shift between the voltage and current. The larger the phase angle, the greater the reactive power generated by the system.

The mathematical relationship among these forms of power can be represented by vectors or expressed using complex numbers:

- S = P + jQ (where j is the imaginary unit).

A vector diagram "power triangle" gives a graphical representation of how all these quantities are related.

> **NOTE: Power engineering**, *also called* **power systems engineering**, *is a subfield of energy engineering and electrical engineering that deals with the generation, transmission, distribution and utilization of electric power and the electrical devices connected to such systems including generators, motors and transformers. Making sure that the voltage, frequency (if AC), and amount of power supplied to a load(s) is in alignment with expectations is of principal importance in power system engineering.*

> **CLARIFICATION:** *Electricity flows readily in some materials but not in others. What differentiates materials is primarily the atomic structure of the matter that comprises them. Some conduct electricity readily; they are of course called conductors. Typical good electrical conductors include copper, aluminum, gold and other metals, and water. Materials that do not conduct electricity are called insulators.*

6.1 Voltage (a.k.a., electric potential difference or electromotive force, EMF)

Voltage is a force/pressure that makes electricity move through a conductor. It is the potential energy source in an electrical circuit that makes things happen. Voltage is also called electric tension and electromotive force (EMF). Voltage (or electric potential) refers to the pressure that pushes electric charges in a circuit. Voltage is the pressure that drives the current. Technically, the voltage is the difference in electric potential between two points. Therefore, voltage is always measured between two points; for example, between the positive and negative ends of a battery, or between a wire and ground.

The voltage, or potential difference from point a to point b is the amount of energy in joules (as a result of electric field) required to move 1 coulomb of positive charge from point a to point b. A negative voltage between points a and b is one in which 1 coulomb of energy is required to move a negative charge from point a to b. If there is a uniform electric field about a charged object, negatively charged objects will be pulled towards higher voltages, and positively charged objects will be pulled towards lower voltages. The potential difference/Voltage between two points is independent of the path taken to get from point a to b. Thus, the voltage from a to b + the voltage from b to c will always equal the voltage from a to c.

Voltage (V) is a measure of the pressure applied to electric charges to make them move. It is a measure of the strength of the current in a circuit and is measured in volts (V). Voltage is the electric power system's potential energy source. Voltage does nothing by itself, but has the potential to do electrical work (i.e., transfer energy). Voltage is a push or a force. The basic unit (measurement) of electromotive force (EMF) is the volt. Voltage is the amount of potential energy that an electron gains or looses by traveling from one potential to another potential. In this way, voltage is very similar to potential energy in kinetics - if I lift a ball, the ball's properties doesn't change but it gains potential energy. Volts means volume.

> **Transformers** *either they "step up" or "step down" voltage.*

There are two types of voltage, DC voltage and AC voltage. The DC voltage (direct current voltage) always has the same polarity (positive or negative), such as in a battery. The AC voltage (alternating current voltage) alternates between positive and negative. For example, the voltage from the wall socket changes polarity 60 times per second (in America). The DC is typically used for electronics and the AC for motors. Normally, voltage is either constant (i.e., direct) or alternating. If voltage is constant (i.e., direct), then current is continuous/direct. If voltage alternates so does current.

> **ANALOGY:** *In water systems, voltage corresponds to the pressure that pushes water through a pipe. The pressure is present even though no water is flowing.*

Voltage is:

- Voltage is a measure of how much energy is delivered to charge.
- Voltage isn't a property of electrons. However, in electronics, charge is generally carried by electrons.
- Voltage unit is potential energy per change: V = potential energy / charge
- Voltage, or electric potential, is the amount of potential energy (joules) that any "charged body" within an electric field will have, for every 1 coulomb of electric charge in it.
- The potential energy does directly translate into kinetic energy if there is only negligible "friction". For example, in an (evacuated) cathode ray tube. The kinetic energy of an electron is indeed measured in "electron volts", eV, the energy (as charge) an electron gains or loses when charges moving through a potential difference of 1 Volt.
- Voltage is a property of an electric field. Note that a gravitational field behaves like an electric field wherein objects are pulled to together. Drop a stone in a gravitational field and it will accelerate downwards, taking energy from the field.
- Using the water analogy, if a tank of water were suspended one meter above the ground with a 1-centimeter pipe coming out of the bottom, the water pressure would be similar to the force of a shower. If the same water tank were suspended 10 meters above the ground, the force of the water would be much greater, possibly enough to hurt you. Just as the 10-meter tank applies greater pressure than the 1-meter tank, a 10-volt power supply (such as a battery) would apply greater pressure than a 1-volt power supply. To remain useful, however, the velocity of the water must be excluded from the analogy. The speed of the flowing water (velocity) increases with pressure, while the speed at which an electric charge propagates through any particular medium is constant even if the "pressure" (voltage) is increased. A poorer analogy might be a tube filled with balls. Apply a force to the ball at one end and it will push the ball at the other end out. Apply a continuous voltage to a wire and the electric charges will move in one continuous direction, "forcing out" the charges at the "positive" end (and entering the power source). The amount of force applied corresponds to the voltage applied to the wire.
- For the water pipe analogy, charge (coulombs) is analogous to the volume of water (gallons), current (amps) is analogous to flow rate of water (gallons per minute), and voltage is analogous to the water pressure that is causing the flow.

The volt is defined as the energy transfer per coulomb of charge as charges move between two points in a circuit.

- $V = \Delta W / \Delta Q$
- Energy change per unit charge (so that 1 V = 1 J C-1)

> **NOTE: Phantom Voltage** *or "***induced voltage***" is the result of wire or other metal components appearing to be energized when they in fact are not. When ungrounded wiring (e.g., Knob & Tube) wiring or older ungrounded romex-type wiring is present, and a metallic pathway (wires and conduit) is added to these old circuits, then the metal wires and/or conduit will pick up an induced voltage merely by being in proximity to the hot conductor in the circuit. The ungrounded wire and conduit, and anything attached to it that is conductive and not grounded, will also appear "energized" (i.e., "hot") when tested with a voltage indicator tool/instrument. Phantom voltage can make the metal sides of ungrounded technological devices like refrigerators, metal light fixtures, metal surface conduits, and metal junction boxes*

appear energized (when they are not actually energized).

6.2 Electrical current (current)

NOTE: *The electrons that move as an electrically charged current come from the conductors (and other connected sources such as a battery, photovoltaic cell, etc.) Remember that in AC, the electrons don't actually travel (as in, DC), they oscillate (as in, AC).*

Because there are two types of voltage, there are also two types of current. There is direct current that moves in one direction, and there alternating current that alternates backward and forward (two directions).

Electric current is the rate of flow of [electrical] charge [carriers]. Whenever electric charges moves or flow, that is called an electric current. The words "electric current" are the same as the words "charge flow." The rate of flow of electric charge is called electric current and is measured in Amperes. Current is the amount of [electric] charge passing through a space per unit time; the rate at which charges flow past a point in space. Current is a physical quantity that can be measured and expressed numerically. If charge is like air, then electric current is like wind. Or, if charge is like water, then electric current is like "volume per second" of water flow. An electric current is the directed movement of electric charge as uni-directional (direct current; DC) or alternating (alternating current; AC). The thing to remember when thinking about energy transfer with AC is that energy still flows from source to sink any time current flows, regardless of the direction of that current. Herein, current density is the electric current per unit area of cross section (amperes per volume).

The term DC is used to refer to power systems that use only one polarity of voltage or current. DC current charge is continuous, while AC current alternates between positive and negative charge. An electric current that flows continuously in a single direction is called a direct current, or DC. The voltage in a direct-current circuit must be constant, or at least relatively constant, to keep the current flowing in a single direction. A sine wave of DC current is a flat line, while AC is an alternating wave of a specific hertz rating. Current in a house circuit (AC) flows in one direction, same as in a battery (DC), hence the polarized plugs and receptacles.

The strength of the current is dependent on the size of the induced charge and the electric resistance the connection. Electromagnetic induction is the generation of voltage across a conductor situated in a changing magnetic field. The difference between the potentials is called the electrical potential difference and is commonly called induced voltage. This generates electric power through the flow of electric charge if both ends of the wire are connected with a conductive wire.

Using the flow of water provides a reasonable analogy for understanding the flow of electric charge. The flow of electric charge in a circuit is similar to water flowing through a hose. If you could look into a hose at a given point, you would see a certain amount of water passing that point each second. The amount of water depends on a number of variables, including how much pressure is being applied (i.e., how hard the water is being pushed). It also depends on the diameter of the hose. Given available water, the harder the pressure and the larger the diameter of the hose, the more water passes each second. The flow of electric charge (measured as electrical current) through a wire depends on the electrical pressure pushing the charges and on the cross-sectional area of the wire.

NOTE: Resistance *is a material's tendency to resist the flow of charge (i.e., the tendency to resist electrical current).*

There are two types of electrical current: alternating current (AC) and direct current (DC). There are two principal types of flowing electric charge (electricity as electric current), and they are used in most cases for very different purposes: direct current (DC) and alternative/ alternating current (AC). Often, electrical power is named after the type of current carrying the electrical power, 'DC power' or 'AC power'. Hence, there are two type types of electrical power: alternating current power (AC power) and direct current power (DC power). Either form can be technically converted into the other form.

CLARIFICATION: *You might be wondering why is it called Alternating current when the voltage is the one that switches from positive to negative. It is called alternating current because as we said above, voltage is the pressure that pushes the current thought the circuit, so if the voltage alternates, the current must also alternate in direction as it is being pushed by the voltage in an opposite direction each time.*

6.3 Two types of electricity

There are two different ways that electricity (electrical power) is produced, and they are used in most cases for very different purposes. They can also be converted from one form to another. The two types of electricity are: direct current (DC) and alternating current (AC). Remember that to produce electricity (electrical power), both current and voltage must be present; hence, when speaking of the presence of electricity (electrical power) it is most appropriate to write 'direct-current (DC) voltage' and 'alternating-current (AC) voltage', although often, only the acronyms DC and AC may be written (they imply the presence of voltage).

6.3.1 Direct-current voltage (a.k.a., DC voltage or DC power)

NOTE: *In a DC system, at the electrical level, it is possible to measure at least: voltage (volts),* *current (amperes), power (watts), and energy (watt-hours or J/s).*

Direct current (DC) is the unidirectional flow of electric charge (electricity flows in one direction). It is the continuous movement of electric charge from an area of negative (-) charge to an area of positive (+) charge -- the difference in charge at two locations connected by a conductor creates an electrical pressure difference (voltage), whereupon charges move from negative to positive until equilibrium is reached. In a direct current system, the voltage does not alternate direction with time. When an electric circuit with DC voltage is complete, the current flows directly, in one direction. It is called direct current as the current is only being pushed in one direction by the voltage. The resultant current creates a unidirectional magnetic field -- a magnetic field with a constant orientation. Hence, a DC current generates a constant magnetic field, and follows the "right hand rule" [wikipedia.org]:

- Induced current 'I' (middle finger)
- Magnetic field 'B' (index finger)
- Motion 'F' (thumb finger)

Direct current is produced by sources such as batteries, power supplies, thermocouples, solar cells, or dynamos (DC generators). Direct current may flow through a conductor such as a wire, but can also flow through semiconductors, insulators, or even through a vacuum, as in electron (dielectric) or ion beams. The electric current flows in a constant direction, distinguishing it from alternating current (AC), which alters its direction of flow at a time interval. Rather than oscillating back and forth as AC does, DC provides a constant uni-directional flow of current.

NOTE: *Direct current was formerly known as galvanic current.*

In a direct current circuit, the power flowing to the load is proportional to the product of the current through the load and the potential drop across the load. Energy flows in one direction from the source to the load. In AC power, the voltage and current both vary approximately sinusoidally.

NOTE: *If the electron flow of a direct current (DC) were converted to a sound, then a DC power signal would sound like a steady tone.*

Almost every electronic device uses DC, and cannot use AC. This is something most people don't realize when they plug some device into a wall outlet. Just because you plug a device into the wall doesn't mean the circuitry inside operates on AC. Very few electronic devices actually can use AC. Almost everything from LED lighting, televisions, stereos, phonographs, tape decks, CD/DVD players, computers, printers, clock radios, battery chargers, along with just about anything that has a micro processor inside or is otherwise computer controlled, all require DC power (if AC power is present, then a connected/internal "power supply" unit must convert AC to DC). If DC were available to begin with there would be no need for AC to DC "power supplies".

NOTE: *The AC to DC power converter for a laptop dissipates (i.e., loses) energy through heat. If the electrical power system was DC, then users could plug in their electronics direct to DC without any need for conversion and any loss of energy to heat. Just like there are standardized plugs for AC, there is likely to be more standardization of DC plugs beyond the relatively low voltage USB DC standard connector.*

A DC voltage source has both negative and positive terminals, and produces a voltage (or potential difference) between those terminals.

Water flow can be used analogistically to describe DC and AC:

1. **Water analogy for DC:** Direct current is like the moving water in a calm river, which has a uniform velocity in and flows only in one direction (from the hills to the sea). DC flows from high potential to low potential, in one direction.
2. **Water analogy for AC:** Alternating current is like the water that is continuously moving forward and backward, such as water waves hitting the beach and receding back. Similarly, AC changes its direction after a particular interval of time.

6.3.2 Alternating-current voltage (a.k.a., AC voltage or AC power)

NOTE: *In an AC system, at the electrical level, it is possible to measure at least: voltage (volts as a function of time, V(t), and as waveform amplitude, Vp), current (amperes), [waveform] frequency (hertz), [waveform] phase (degrees), power (watts), and energy (watt-hours or J/s).*

Alternating current (AC) is a flow of electric charge that periodically reverses/changes direction - an alternating/changing flow of electric charge (i.e., the current *alternates* its direction over time as it flows; electricity flows in both directions). It is essentially a steady sine wave and the reversal of polarity that, over time, generates an oscillation frequency (i.e., a back and forth movement of electric charge at a "speed" known as a 'frequency', which is measured in the unit, hertz). Alternating currents are accompanied (or caused) by alternating voltages. AC as its definition is alternating current (i.e the amplitude of the current is different for different instants of time). Alternating current is not a single, constant voltage, but rather as a sinusoidal wave that starts at zero, and over time, increases to a maximum value, then decreases to a minimum value, and repeats. AC always implies alternate (and therefore varying). AC switches polarity over time, in a precise sinusoidal-like manner, causing electric charges ("electrons") to pulse back and forth [over a material known as a "conductor"]. Alternating voltage pushes and pulls the charge backwards and forwards in the conductor (e.g., a wire). The two ends of the circuits become both the positive and negative pole at different times. In its pulsating movement, alternating current creates a moving magnetic field inside the conductor.

NOTE: *In an alternating current (AC) system there is no static "plus" (+) or "minus" (-) in the circuit, because each side (terminal) of the circuit is a (+) and (-) at different times.*

The two principal properties of an AC electrical current are: voltage and frequency. These two properties differ between market-State regions. A voltage of (nominally) 230V and a frequency of 50Hz is used in Europe, most of Africa, most of Asia, most of South America and Australia. In North America, the most common combination is 120V and a frequency of 60Hz. Other voltages exist, and some countries may have, for example, 230V and 60Hz.

In AC power system, the (power) "line frequency" or "mains frequency" is the frequency of the oscillations/cycles of alternating current (AC current cycle). Essentially, frequency refers to how often the current changes direction. The two principal frequencies used throughout the world form common (non-specialized) AC power are 50Hz and 60Hz (50-cycle and 60-cycle). For instance, a generator with one pole (one alternating current cycle per revolution) turning at 3600 rpm will rotate 60 times in one second, thus generating 60 alternating current cycles per second or 60 Hz current. It follows that a machine rotating at 1800 rpm will require four poles to produce the same 60Hz current.

Electrical generators (AC) generally seek to produce electric power where the voltage waveform has only one frequency associated with it, the fundamental frequency (e.g., 50Hz or 60Hz). Hence, when an electrical circuit is connected to the coils of an operating generator, there will exist an oscillating electric current "surging" back and forth through each coil at a rate of #Hz times a second (e.g., 50Hz or 60Hz).

NOTE: *The sequence of successive peaks of the currents (i.e., phases) causes a magnetic field to form and move around the stator air gap.*

The appearance of additional frequencies (frequency waveforms) produces harmonics. Harmonics are distortions of the pure sinusoidal waveform. It is the sinusoidal component of a periodic wave or quantity having a frequency that is an integral multiple of the fundamental frequency. Harmonics are a multiple of the fundamental frequency. Some references refer to "clean" or "pure" power as those without any harmonics. Some loads cause the voltage and current waveforms to lose this pure sine wave appearance and become distorted. Note that in acoustics (as in, music), harmonics are called overtones.

NOTE: *"Mains electricity" (mains AC) is the general-purpose alternating-current (AC) electric power supply fed into commercial buildings and residential homes by the electric power industry in early 21st century society. Mains electricity is also referred to by several names including household power, household electricity, house current, powerline, domestic power, wall power, line power, AC power, city power, street power, and grid power.*

Frequencies can vary from as low as 10Hz (or less) to as high as 400Hz (or more) for specialized AC power systems. Several factors influence the choice of frequency in an AC system, and the design of generators, transformers, transmission lines, and end-load devices depend on the power frequency. The usage of AC technology of a different frequency rating than the one specified by the manufacturer of a device can be dangerous.

NOTE: *AC can be radiated from an antenna, and this capability is responsible for radio communication (e.g., radio, wifi, bluetooth, cellular communication, etc.). In other words, alternating electromotive force (in the conductor also radiates radio frequencies (as electromagnetic frequency waves, EMF -- not the same as electromotive force EMF) from the conductor. This electromagnetic frequency field (EMF) reverses its polarity when it moves under magnetic poles of opposite polarity.*

The simplest form of AC power consists of a source and a linear load, and both the current and voltage are sinusoidal.

1. If the load is <u>purely resistive</u>, the two quantities reverse their polarity at the same time. At every instant the product of voltage and current is positive or zero, with the result that the direction of energy flow does not reverse. In this case, only active power is transferred.
2. If the loads are <u>purely reactive</u>, then the voltage and current are 90 degrees out of phase. For half of each cycle, the product of voltage and current is positive, but on the other half of the cycle, the product is negative, indicating that on average, exactly as much energy flows toward the load as flows back. There is no net energy flow over one cycle. In this case, only reactive power flows—there is no net transfer of energy to the load.
3. <u>Practical loads</u> have resistance, inductance, and capacitance, so both active and reactive power will flow to real loads. In an AC system, power is measured as the magnitude of the vector sum of active and reactive power. **Apparent power** is the product of the root-mean-square of voltage and

current. In alternating current (AC) circuits, energy storage elements such as inductors and capacitors may result in periodic reversals of the direction of energy flow leading to the creation of active power and reactive power

A. The portion of power that, averaged over a complete cycle of the AC waveform, results in net transfer of energy in one direction is known as **active power** (sometimes also called **real power**).
B. The portion of power due to stored energy, which returns to the source in each cycle, is known as **reactive power**. Energy stored in capacitive or inductive elements of an AC power network give rise to reactive power flow. Reactive power flow strongly influences the voltage levels across the network. Voltage levels and reactive power flow must be carefully controlled to allow a power system to be operated within acceptable limits.

When there is inductance or capacitance in the circuit, the voltage and current waveforms do not line up perfectly. The power flow has two components - one component flows from source to load and can perform work at the load, the other portion, known as "reactive power" is due to the delay between voltage and current, known as phase angle, and cannot do useful work at the load. It can be thought of as current that is arriving at the wrong time (too late or too early). To distinguish reactive power from active power, it is measured in units of "volt-amperes reactive" or var. These units can simplify to Watts, but are left as VAR to denote that they represent no actual work output.

Alternating current has the compelling advantage over direct current; its voltage can be changed easily and efficiently by a transformer. A transformer is generally composed of a closed iron core surrounded by two windings (first and fourth principles of electrical machinery). The ratio of the voltages in the two windings is the same as the ratio of the number of turns, and the ratio of the currents inversely, so that the power remains the same. The ratio of the voltage fluctuations on each side is the ratio of the number of turns the wire makes around the core (on each side). Since there are no mechanical parts, the efficiency of transformers is very high, and maintenance very low. Alternating current is transformed to higher voltage and smaller current for transmission, and back to lower voltages for use. Transformers with taps can be used to obtain a series of voltages if desired. In fact, an almost continuous voltage variation without loss is possible. It is not easy to change DC voltages. One way to do this was to use a dynamotor, which had a normal field winding, but dual armature windings and two commutators. One winding was supplied at the input voltage and drove the dynamotor by motor action. The other winding supplied the output voltage. This can really be considered a kind of AC transformer. The input commutator creates AC from DC, and the output commutator changes the new AC voltage to DC.

1. Devices that can be designed to work effectively with AC: electric light-bulbs; electric heating elements; small electric motors (food mixers and vacuum cleaners).
2. Devices that preferably or necessarily run on DC: large electric motors; electronics.

NOTE: *Earth has a magnetic field, and it is moving in a more or less circular path around the moon, and so, one could theorize that a conductor on the moon may have such a wave induced into it. However, there is not any knowledge, presently, of any organisms in nature that use a reversal of polarity as a force in their sustainment.*

6.4 Electrical charge

NOTE: *Electricity is commonly defined in practical application (non-theoretical scientific study) as the flow of electric charge (electric current).*

The material things around us are made of 'atoms'. Atoms are the fundamental building blocks of all molecules, and they consist of three types of 'particles': protons, neutrons, and electrons. Of these three subatomic particle types, two (protons and electrons) carry a net electric charge, while neutrons are neutral and have no net charge. Atoms have a "positively" charged nucleus (containing the protons with a "positive" charge and the neutrons with no net charge, hence "positively" charged). The nucleus is surrounded by "negatively" charged electrons.

NOTE: *If an atom has an equal number of protons and electrons, its net charge is 0. If it gains an extra electron, it becomes negatively charged and is known as an anion. If it loses an electron, it becomes positively charged and is known as a cation.*

Unlike protons, electrons can move from atom to atom. Hence, electrons are considered mobile charges (i.e., they are the mobile charge carriers in an electric circuit). In physics, 'electrons' are the smallest unit of "negative" electric charge, and protons are the smallest unit of "positive" electric charge. Summarily, there are two types of electric charge: "positive" (proton) and "negative" (electron), with the neutrons having a neutral (0) charge. It is possible to encounter free positive charges (e.g. a free proton or ion) in atomic or nuclear physics, or in chemistry. There are also positively charged electrons (positrons), but they occur under special conditions and do not survive long.

> **NOTE:** *If there is a quantity of charge, it cannot be destroyed, it can only be moved from place to place.*

The electric charge (elementary charge) is one of the fundamental quantities/constants of physics, along with mass and time. An elementary charge -- that of a proton or electron -- is approximately equal to 1.6×10^{-19} Coulombs.

The motion of charge carriers is electric 'current'. In other words, when charges move they form a "flow of electric charge", which is called an electric current. In electricity and electronics, the negative charges are the electrons, and the currents almost always refer to the movement of electrons. Note here that the direction of the electric current is always opposite the motion of the electrons, because someone in the past decided that the direction of the current should be in the direction of 'positive' charges, and scientists have not updated their language since.

> **NOTE: Electrostatics** *is a branch of physics that deals with the phenomena and properties of stationary or slow-moving electric charges.* **Electrodynamics** *is a branch of physics that deals with the phenomena and properties of moving charges.*

In specific, an **electric charge** is the physical property of matter that causes it to experience a force when placed in an electromagnetic field; this force is known as the 'electric force'. Charges produce electromagnetic fields, which act on other charges. Electrically charged matter is influenced by, and produces, electromagnetic fields (EMF). The interaction between a moving charge and an electromagnetic field is the source of electromagnetic force, which is one of the four fundamental forces.

The concept 'electrical energy' refers to energy carried by [moving] electrical charges. Note here that an electrical charge is not energy; it carries energy. The faster electrical charges are moving, the more electrical energy they carry. Note here that when electrical charges are moving (current), they are considered a form of kinetic energy. Whereas a static charge (unmoving charge) contains potential energy, and when it moves, this energy is said to be "converted" to kinetic energy.

> **NOTE:** *Charge carriers are "pushed" around a circuit by an electromotive force (EMF or voltage). Despite its name, EMF is not a force but a voltage, measured in volts. In other words, the pressure that moves charge carriers around an electric circuit, and thus, transfers energy from source to load, is called an electromotive force (EMF), which is not a force, but is in fact voltage. The electromotive force is voltage across a source of electrical energy, and therein, potential difference is voltage across a component that uses electrical energy. EMF is energy supplied per coulomb. In other words, the volt is defined as the energy transfer per coulomb of charge as charges move between two points in a circuit:* $V = \Delta W / \Delta Q$. *Charge carriers (electrons being one of such) can be used to transmit an electromotive force (usually called just voltage).*

If a current is present, then there is a net motion of charge carriers, and "electrical" energy is being transferred. However, it is generally not correct to say that an electric current is "a flow of electrons"; instead, it is more correct to say that electric current is a flow of electric charge. Charge can be positive (protons) or negative (electrons), and both types of charged particles can and do flow in electric circuits. In different media, different particles serve to carry charge:

1. In metals - the charge carriers are electrons.
2. In electrolytes (e.g., salt water) - the charge carriers are ions, and atoms or molecules that have gained or lost electrons so they are electrically charged. Atoms that have gained electrons so they are negatively charged are called anions, atoms that have lost electrons so they are positively charged are called cations.
3. In plasma - the electrons and cations (cat-ions) of ionized gas act as charge carriers.
4. In a vacuum [tube] - free electrons can act as charge carriers.
5. In a semiconductor (e.g., transistor) - electrons and traveling vacancies in the valence-band electron population (called "holes") are the charge carriers.
6. In hydrogen fuel cells and water ice - current consists of a flow of protons, which are the charge carriers.

In physics, it is presently understood that the "electrons" do physically move (when a voltage is applied) both in AC and DC, though slowly in DC. In a DC circuit, the electrons move in one direction. In an AC circuit, the electrons don't move continuously forward; instead, they move backwards and forwards (i.e., they oscillate), and may be said to "vibrate". The thing to remember when thinking about energy transfer via electric current, regardless of DC or AC, is that energy flows from source to load any time electric current flows, regardless of the direction of the electric current (or moving "electrons").

> **CLARIFICATION:** *There are [at least] two things moving through an electrical power system: "electrical" energy and electric charge carried by a charge carrier (e.g., electron).*

Here, the term "charging" refers to giving an object/system a[n electric] charge. There are three common methods of "charging":

1. Charging by friction - Rubbing two different materials together, a process known as charging by friction (a.k.a. charging by rubbing), is the simplest

way to give something a charge.

2. Charging by induction - It is possible to charge a conductor without touching it. Charging by induction requires a procedure involving [at least] two objects and a ground connection.
3. Charging by conduction - The two objects will come into actual physical contact with each other, and contact transfers the charge (this is why it is sometimes called "charging by contact").

6.5 *Electrical power systems*

An electrical power system is a network of electrical components interconnected to supply, transfer and distribute, and use, electrical power.

> **NOTE:** *In order to function, an electrical power system must form an electrical circuit.*

In general, when electrical power is supplied and used over a land area larger than a single building, then the power system is known as "the electrical grid", which can be broadly divided into the generators that supply the power, the transmission system that carries the power from the generating source(s) to the load(s), and the distribution system that feeds the power to end-point uses. Hence, a "wide-scale" electrical power system consists of a power station(s) connected to a transmission and distribution system. A localized electrical power system may consist only of an electrical power generator and distribution system. In general, electrical power systems also involve electrical power storage and recovery systems. Herein, the electrical transmission grid is an electrical circuit.

All electrical power systems are composed of the following:

1. **The supplying source:** All power systems have one or more sources of power. For some power systems, the source of power is external to the system, and for others it is part of the system itself. There are two principal types of power supply: alternating current (AC) and direct current (DC).
2. **The electrical circuit with an electrical load:** Power systems transfer energy to loads that perform a function. In general, it is a composition of material converts electrical energy to some other form of useful energy. Most loads expect a certain voltage, and for alternating current devices they necessitate a certain frequency and number of phases. Every load will have a wattage, which specifies the amount of electrical power the load consumes. At any one time, the net amount of power consumed by the loads on a power system must equal the net amount of power produced by the supplies, less the power lost in transmission.
3. **Conductors:** Conductors carry power from the generator to the load, or from a generator to an induction point, and then, from the opposite side of the induction point to the load. In a grid, conductors may be classified as belonging to the transmission system, which carries large amounts of power at high voltages (typically more than 69 kV) from the generating centres to the load centres, or the distribution system, which feeds smaller amounts of power at lower voltages (typically less than 69 kV) from the load centres to nearby homes and industry. There are also conductors within electronic devices themselves. Note that different materials (and different combinations of materials) carry different conductivity measures. Conductors are typically rated for the maximum current that they can carry at a given temperature rise over ambient conditions. As electrical current flow increases through a conductor it heats up. For 'insulated conductors', the rating is determined by the insulation, and for 'overhead conductors', the rating is determined by the point at which the sag of the conductors would become unacceptable. Electric conductors are substances that possess the quality of electric charge. Within all metals there is a substance which can move.
4. **Power electronics**: Power electronics are semi-conductor based devices that are able to switch quantities of power ranging from a few hundred watts to several hundred megawatts. The classic function of power electronics is rectification, or the conversion of AC-to-DC power, power electronics are therefore found in almost every digital device that is supplied from an AC source either as an adapter that plugs into the wall (see photo in Basics of Electric Power section) or as component internal to the device. High-powered power electronics can also be used to convert AC power to DC power for long distance transmission in a system known as HVDC. HVDC is used because it proves to be more economical than similar high voltage AC systems for very long distances (hundreds to thousands of kilometres).
5. **Earthing system** (**grounding system**) is circuitry which connects parts of the electric circuit with the ground (earth), thus defining the electric potential of the conductors relative to the Earth's conductive surface. The purpose of an earthing system is to provide an alternative path for the fault current to flow so that it will not endanger the user, ensure that all exposed conductive parts do not reach a dangerous potential, maintain the voltage at any part of an electrical system at a known value

and prevent over current or excessive voltage on the appliances or equipment. Different earthing systems are capable of carrying different amounts of over current. There are two types of earthing systems:

A. Unearthed system: IT system
B. Earthed system: TT system; TN system (TN-S, TN-C, TN-C-S)

6. **Protective devices:** Power systems contain protective devices to prevent injury or damage during failures.
   A. The most common 'protective device' is a fuse. Fuses must be replaced as they cannot be reset once used (i.e., blown). Also, fuses can be inconvenient if the fuse is at a remote site or a spare fuse is not available.
   B. Circuit breakers are devices that can be reset after they have broken current flow.
   C. Protective relays are used in high power applications. They detect a fault and initiate a trip.
   D. Enclosing an arc chamber and flooding it with sulfur hexafluoride (SF6), a non-toxic gas that has sound arc-quenching properties.
   E. Residual current devices (RCDs) - In any properly functioning electrical appliance the current flowing into the appliance on the active line should equal the current flowing out of the appliance on the neutral line. A residual current device works by monitoring the active and neutral lines and tripping the active line if it notices a difference. Residual current devices require a separate neutral line for each phase and to be able to trip within a time frame before harm occurs.
7. **Supervisory Control And Data Acquisition (SCADA)** is used in large electric power systems for tasks such as switching on generators, controlling generator output and switching in or out system elements for maintenance.

In the market, electrical power systems are generally subdivided into residential power systems (small scale) and commercial power systems (large scale):

1. Residential dwellings almost always take supply from the low voltage distribution lines or cables that run past the dwelling. These operate at voltages of between 110 and 260 volts (phase-to-earth) depending upon national standards. Each dwelling has its own circuit breaker.
2. Commercial power systems are uniquely designed for load flow, short-circuit fault levels, and voltage drop for steady-state loads and during starting of large motors. Typically one of the largest appliances connected to a commercial power system is the HVAC unit.

6.5.1 The electric circuit

> **NOTE:** *Power in an electric circuit is the rate of flow of energy past a given point of the circuit.*

An electric circuit is a path in which electrons from a voltage or current source flow. The point where those electrons enter an electrical circuit is called the "source" of electrons. The point where the electrons leave an electrical circuit is called the "return" or "earth ground". An electrical circuit is a network consisting of a closed loop, giving a return path for the current. An electrical circuit is a path or line through which an electrical current flows. The path may be closed (joined at both ends), making it a loop. A closed circuit makes electrical current flow possible. It may also be an open circuit where the electron flow is cut short because the path is broken. An open circuit does not allow electrical current to flow. Hence, a working **circuit** is a closed loop -- to be a circuit, all charge must find a path back to its source, regardless of the source (including a battery or a transformer on the pole; this is part of "Kirchhoff's current law").

> **NOTE** *that in the case of static electricity, the "Kirchhoff current law" will accept a temporary delay and storage of charge. The most commonly encountered real world example of this is on a cold and dry day: you walk across the carpet and touch a doorknob and experience a spark. As you walked, your socks picked up charge from the carpet; it flowed out onto your body, and when you touched the doorknob that charge began its journey back to the carpet fibers from where it came via material from which the door is composed.*

Electric power is transferred to other carriers of energy when electric charges move through loads in electric circuits. From the standpoint of electric power, the components in an electric circuit can be divided into two categories:

1. **Active devices** or **power sources:** When electric charges move through a potential difference from a higher to a lower voltage, that is when conventional current (positive charge) moves from the positive (+) terminal to the negative (−) terminal, work is done by the charges on the device. The potential energy of the charges due to the voltage between the terminals is converted to kinetic energy in the device. These devices are called passive components or loads; they "consume" electric power from the circuit, converting it to other forms of energy such as mechanical work, heat, light, etc. In alternating current (AC) circuits the direction of the voltage periodically reverses, but the current always flows from the higher potential to the lower

potential side

2. **Passive devices** or **loads:** If the charges are moved by an 'exterior force' through the device in the direction from the lower electric potential to the higher, (so positive charge moves from the negative to the positive terminal), work will be done on the charges, and energy is being converted to electric potential energy from some other type of energy (e.g., mechanical energy or chemical energy). Devices in which this occurs are called active devices or **power sources**; such as electric generators and batteries.

Note that some circuitry devices (i.e., devices connected to a circuit) can be either a source or a load, depending on the voltage and current passing through them. For example, a rechargeable battery acts as a source when it provides power to a circuit, but as a load when it is connected to a battery charger and is being recharged.

The three main circuit components are:

1. The resistor
2. The capacitor
3. The inductor

Hence, the types of electrical circuits associated with electrical power production or power conversion systems are:

1. Resistive
2. Capacitive
3. Inductive

Most systems have some combination of each of these three circuit types. These circuit elements are also called loads. A load is a part of a circuit that converts one type of energy into another type. A resistive load converts electrical energy into heat energy.

6.6 Electric[al] power generation (electrical power source)

- Thermal > electric (e.g., thermopile, thermoelectric generator)
- Mechanical > electrical (e.g., alternator)
- Fluidal > electrical (e.g., turbine-electric)
- Chemical > electrical (e.g., chemical battery)
- Magnetic > electrical (e.g., electrical generator)
- Electromagnetic > electrical (e.g., antenna receiver, solar panel)

Electrical power is present when electric charges move (current) through an electric potential difference (voltage). Therein, electrical power is the product of the current and the voltage. Hence, to generate electrical power, a method and/or system must induce voltage across a conductor to produce a current (simplistically, it must produce both voltage (V) and current (I), where Power = VI).

Technically, current is the result of the generation of voltage (electrical potential difference, electromotive force) between two points on a conductive circuit. The product of voltage and current is power, and when power is present, energy is being transferred.

> **DEFINITION:** *A* **power station***, also referred to as a generating station, power plant, powerhouse, or generating plant, is a technical production space (or, location or facility) for the generation of power (generally, electrical power, but if the context is not clear, then it could be any type of power).*

Electric power generation is the process of generating electric power (voltage and current) from other carriers (sources) of energy. Herein, the term 'direct electric power generation' refers to energy transfer methods and technologies that are capable of directly producing electrical power (electricity) from some other type of energy carrying input. There are several fundamental scientific effects that may be applied methodically (procedurally) for producing and/or generating electrical power from other sources/carriers of energy:

1. **Electrostatic effect** (static electricity through friction) - technically, not a form of electric power, but a form of voltage as an imbalance of electric charges within or on the surface of a material. The charge remains until it is forced to move by means of an electric current or electrical discharge. Static electricity is the physical separation and transport of charge (e.g., triboelectric effect, lightning, and friction sparks). The discharge or other movement of static electricity carries electrical power.
   - When an object with a normally neutral charge loses electrons, due to friction, and comes in contact with another object having a normal charge, an electric charge is exerted between the two objects.
2. **Electromagnetic induction effect** (magnetic induction, electromagnetism) - is the production of an voltage (electromotive force) across an electrical conductor due to the conductor's dynamic interaction with a magnetic field. The effect of time-varying magnet fields is to produce a time-varying electromotive force (EMF) that drives/forces charged particles around a circuit. The effect is described exactly by Maxwell's equations. And, electromagnetic induction is based on Faraday's law. Today, electromagnetic induction is the most widely used method for generating electricity. This method generally involves an input of rotational (angular) mechanical energy (e.g., turbine).
   - Induction, as the movement of a conductor

in a magnetic field, directly creates an electric potential (a dynamic time-varying electric field) in the conductor, whereupon charges move, and hence, current flows.
- **Magnetohydrodynamics** (MHD; also magneto fluid dynamics or hydromagnetics) - an electrically conductive fluid passes through a magnetic field, whereupon electrical power may be generated based on Faraday's law of induction. A generator using this mechanism is also known as a magnetohydrodynamics generator.

3. **Electrochemical effect** (electrochemistry) - the direct transfer of chemical energy into electrical energy through chemical reaction (e.g., battery, fuel cell, or nerve impulse). These chemical reactions involve electric charges moving between electrodes and an electrolyte (or ionic species in a solution). Thus, electrochemistry deals with the interaction between electrical energy and chemical change.
   - A chemical reaction in a system directly create an electric potential and current.
   - Combining chemicals with certain metals causes a chemical reaction that transfers electrons.
4. **Photoelectric effect** (**photovoltaic effect**) - a transfer of electromagnetic energy (light) into electrical energy (e.g., photovoltaic solar cells). The photo-electric effect/principle states that a system can only collect/assimilate light when electrons are present. The greater the presence of electrons, the more energy carried by light can be transferred into an electrical circuit.
   - Light (electromagnetic energy) contacting a system directly creates an electric potential and current.
   - Dislodging of electrons from their orbits by light beams creates positively-charged objects.
5. **Thermoelectric effect** (heat) - the direct conversion of temperature differences into electric voltage, and vice versa. A thermoelectric device creates voltage when there is a different temperature on each side (e.g., thermocouples, thermopiles, and thermionic converters). Note, power generation methods that use heat as a primary input are subject to Carnot efficiency limitations. Any thermodynamic driving force (heat) can directly generate electricity.
   - Heat (thermal energy transfer) into a system directly creates electric potential and current.
   - Heating two joined dissimilar materials will cause a transfer of electrons between the materials setting up a current flow.
6. **Thermionic emission effect** (thermal electron emission , Edison effect)- the thermally induced flow of charge carriers from a surface or over a potential-energy barrier. This occurs because the thermal energy given to the carrier overcomes the work function of the material. A thermionic energy "converter" is a device consisting of two electrodes placed near one another in a vacuum. One electrode is normally called the cathode, or emitter, and the other is called the anode, or plate. At a sufficiently high temperature, a considerable number of electrons are able to "escape" the cathode. The electrons that have escaped from the hot cathode form a cloud of negative charges near it called a space charge. If the plate is maintained positive with respect to the cathode by a battery, the electrons in the cloud are attracted to it. As long as the potential difference between the electrodes is maintained, there will be a steady current flow from the cathode to the plate.
   - Freeing electrons from a hot surface causes electrons to "escape".
7. **Piezoelectric effect** (mechanical pressure) - the electric charge that accumulates in certain solid materials (such as crystals, certain ceramics, and biological matter such as bone, DNA and various proteins) in response to applied mechanical stress. The charge comes from the mechanical strain of electrically anisotropic molecules or crystals.
   - Mechanical stress in a system directly creates electric potential and current.
   - Bending or twisting certain materials will cause electrons to drive out of orbit in the direction of the force. When the force is released, the electrons return to their original orbit.
8. **Nuclear transformation** - charged particles (e.g., betavoltaics or alpha particle emission).
   - The "decay" of nuclear potential energy creates electric potential and current in a system.

Electric power generation considerations:

1. How much electrical power does the generator and or generation station need to supply?
2. Is there a proximity requirement between the generator and load/user?
3. What source of primary energy is available in proximity?
4. What is an acceptable length of time for starting the generator (some generators can take hours to start)?
5. Is the availability of the power source acceptable (some sources create safety issues, such as pollution, and some sources are available only periodically such as sunlight and wind)?
6. How should the generator start (some turbines act like a motor to bring themselves up to speed

in which case they need an appropriate starting circuit)?

7. Which of the two types of current-voltage (DC or AC) is required as input?
8. What is the mechanical speed of operation for the turbine, and consequently, what are the number of poles required?

6.6.1 Electromagnetic induction

Generally, electromagnetic induction generators require a source of rotational (angular) mechanical power to produce a dynamic interaction between the conductor and the magnetic field. Anything that produces torque at a sufficient number of rotations per minute (e.g., turbines, motors and engines) can be used (as a mechanical power source) to turn the rotor within an electromagnetic induction generator to produces electric power.

Presently, most electrical power on Earth is generated by electromagnetic induction through turbine-generator systems. The turbine produces mechanical shaft power as the power input into an electromagnetic induction generator, which outputs electrical power (AC only, not DC). The turbine turns the rotor of the generator with mechanical torque, and the generator transfers this “mechanical” energy to “electrical” energy.

AC power may be directly generated through rotating electromagnetic equipment. An AC power generator is made up of a stator (which remains still), a rotor (which rotates), and electromagnetic fields that generate electromotive force (i.e., voltage) in conductors. When the rotor turns, the magnetic field begins moving in relation to a conductor. Whereupon, voltage is induced in the conductor (coil), which can be measured between the two ends of the coil. The two ends of the coil are called the generator’s terminals. An electrical circuit is hooked up to these terminals, and the ensuing electric power is transmitted to the load on the circuit.

Rotating electromagnetic generators directly produce AC current, which can be modified to produce DC current. There are two types of rotating electromagnetic (AC) generator: synchronous and asynchronous.

6.6.2 Dispatchability

Electrical power generating systems may be classified according to their dispatchability (duty). Dispatchable generation refers to sources of electricity that can be dispatched at the request of power grid users (and hence, intersystems team operators). In other words, dispatchability refers to the ability of generating plants to be turned on or off, or can adjust their power output, accordingly to demand, and hence, the output of a control system [as part of the Decision System]. Power stations are either dispatched (scheduled), or non-dispatched (non-scheduled):

1. **Dispatched: Base load power plants (load matching)** run nearly continually to provide that component of system load that doesn’t vary during a day or daily cycle (e.g., week). Baseload plants can be highly optimized for low fuel cost, but may not start or stop quickly during changes in system load. Examples of base-load plants would include large coal-fired and nuclear generating stations, or hydro plants with a predictable supply of water.
2. **Dispatched: Peaking power plants (peak matching)** meet the daily peak load, which may only be for one or two hours each day. They operate in tandem with base load power plants as required to ensure production capacity of the system during load peaks. Peaking plants include simple cycle gas turbines and sometimes reciprocating internal combustion engines, which can be started up rapidly when system peaks are predicted. Hydroelectric plants may also be designed for peaking use.
3. **Dispatched: Load following power plants (lead-in times)** follow the variations in the daily and daily cycle (week), at a lower resource usage than peaking plants, and with more flexibility than baseload plants.
4. **Dispatched: Backup for base-load generators** - Nuclear power plants, for example, are equipped with nuclear reactor safety systems that can stop the generation of electricity in less than a second in case of emergency.
5. **Non-dispatchable plants (intermittent renewable energy, frequency regulation)** involve quality and stability changes in the electricity output sent into the system because of a change in the frequency of electricity transmitted; renewable sources such as wind and solar are intermittent and may need flexible power sources to smooth out their changes. While their long-term contribution to system energy supply is predictable, on a short-term (daily or hourly) base it may vary from predictable for some sources (e.g., solar) to unpredictable [given present knowledge] for other sources (e.g., wind). In some cases their generated power can be deferred (e.g., to a battery), and in other cases it must be used as it is generated.

NOTE: *The division between dispatchable power plants and renewable power plants is something of a false dichotomy, because if you have sufficient wind or solar generators, then they can be turn on and off (e.g., number of wind turbines operating) and adjusted (e.g., angle of incidence for solar) to control their power output. A solar system obviously cannot be dispatched at night, but that is predictable. Cloud cover is not as yet as predictable. In a planned and integrated* *power system there are known variables and parameters attached to different power sources: predictability and variability therein, startup and shutdown timing, and output adjustability.*

Electrical grid dispatchable variables include:

1. Electrical system balancing - changes in power demand require changes in supply in order to ensure load following and frequency control
   - Load matching - slow changes in power demand require changes in supply in order to ensure balance.
   - Load following - medium (not slow and not short) changes in power demand require changes in supply in order to ensure balance
   - Peak matching - short/rapid changes in power demand require changes in supply in order to ensure balance.
2. Lead-in times - periods during which an alternative source is employed to supplement the lead time required by primary power sources.
3. Frequency regulation or intermittent power sources - changes in the electric power output sent into the system may change quality and stability of the transmission system itself, because of a change in the frequency of electricity transmitted.
4. Backup for base-load generators - safety systems that can stop the generation of electricity in less than a second in case of emergency.

6.6.3 Electromechanical systems (a.k.a., electrical machines)

NOTE: Electromechanics *combines electrical and mechanical processes and procedures drawn from electrical engineering and mechanical engineering. Electromechanical machines are also known as electrical machines. Here, the word machine added to electrical implies a mechanical element connected to an electrical element.*

The terms 'electromechanical systems' (a.k.a., electromechanical machines/devices) and 'electrical machines' mean the same thing. These terms refer to a machine (mechanical system) combined with an electrical system (electrical device) to transfer (i.e., "convert") mechanical energy to electrical energy, and vice versa. The electromechanical "conversion" process involves the transfer of energy between electrical and mechanical systems (and vice versa), via motion in combination with electromagnetic phenomena (i.e., the electric and magnetic fields). Electrical machinery are devices that utilize electromagnetic phenomena in the transfer of mechanical power into electrical power, and vice-versa. By the classical definition, an electric machine is synonymous with electric motors and electric generators, all of which are electro-mechanical energy "converters". They convert electricity to mechanical power (i.e., electric motor) or mechanical power to electricity (i.e., electric generator).

NOTE: *The process of electromechanical energy transfer ("conversion") is reversible in nature, apart from the losses taking place in the device.*

Herein, the energy is not created or destroyed, but it is transferred between electrical and mechanical carriers. In practice, there are four basic types of electromechanical system (a.k.a., electrical machines):

1. **Transducers**: These electronic devices transfer (or "convert") energy signals from one carrier (or "form") to another. These devices mostly operate on vibrating motion. Examples are microphones, pickups, and speakers.
2. **Mechanical force producers**: These types of devices produce mechanical force or torque based on translatory motion. These devices handle larger energy signals than transducers. Examples are relays, solenoids (linear actuators), and electromagnets.
3. **Continuous electrical power "converters" (rotating electromagnetic machines)**: There are two types of continuous energy "conversion" devices, both of which operate in rotating mode, and are thus referred to as <u>rotating electromagnetic systems</u> (machines, devices, or equipment):
   - **Generators** transfer mechanical energy to electrical energy.
   - **Motors** transfer electrical energy to mechanical energy.
     - Electric generators and motors operate by virtue of induced electromotive force (emf, voltage). The induction of emf is based on Faraday's law of electromagnetic induction. Every generator and motor is a rotating electromagnetic system and has a stator (which remains stationary) and rotor (which rotates).
     - There are three types rotating electromagnetic/ electromechanical machine:
   - **DC machines** - produce DC current (DC generator) or accept DC current (motor)
   - **Asynchronous AC machines** - produce (generator) or accept (motor) AC current in an asynchronous manner. **Induction machines** denote asynchronous machines of which only one winding is energized.
   - **Synchronous AC machines** - produce (generator) or accept (motor) AC current in a synchronous manner.

4. **Transformers:** These devices do not transfer (i.e., “convert”) between mechanical and electrical carriers, but they convert AC current from one voltage level to another voltage level. Although transformers do not contain any moving parts, they are also included in the family of electric machines because they utilise electromagnetic phenomena. Besides transformers, electromagnetic machines link an electrical energy system to a mechanical energy system.

6.6.4 Electrical current for electromechanical systems

An electrical current creates a surrounding magnetic field that is strengthened by passing through an iron core. This principle can be called “electromagnet action.”

1. An electrical current causes a magnetic field that surrounds it like a continuous vortexing torus tube. This field, which is not material, is a region of influence on other electrical currents, magnetic fields, and light (EMR). The field is guided and strengthened by passing through iron. When the current reverses in direction, so does the magnetic field. For instance, electrical currents within the earth [in part] cause its magnetic field. The field acts on a compass needle, which is a magnet (magnetized iron). Compass needles are made of iron alloys which can hold their magnetism for a long time. Conversely, the pointers on mechanical watches are not made of iron. In fact they can be made out of almost anything, so long as it is not iron. (Usually brass, drawn very fine, which can maintain its stiffness while being light enough to be moved easily by the delicate forces employed in watches.) No - one wants their watch to be affected by any magnets bought near the watch, so watches are made from materials that are not attracted to magnets.
2. A force is exerted on an electrical current in a magnetic field perpendicular to the plane of the magnetic field and the electrical current (current is ‘x’ axis, force is ‘y’ axis, and magnetic field is ‘z’ axis). An electrical current in a magnetic field (produced by some other source or electrical current) experiences a force perpendicular to both the direction of the current and the direction of the magnetic field, and reverses if either of these reverse in direction. The force is proportional to the current and to the strength of the magnetic field. This principle can be called “motor action”.
3. A voltage (electromotive force) is induced in a conductor moved in a magnetic field. Note that the voltage is opposite to the electrical current causing a force in the direction of motion by principle 2. This principle can be called “generator action”.An electrical conductor, such as a copper wire, moving in a magnetic field has an electrical current induced in it. This is expressed by the creation of an electromotive force (EMF, measured in voltage), which causes current to flow just like the voltage of a battery connected to a circuit. The effect is maximum when the wire, the motion, and the magnetic field are all mutually perpendicular. Electromotive force (EMF) is the voltage generated by a source like battery or generator. Voltage can be measured between any two points, but EMF exists only between the two ends of a source. Voltages in a circuit called ‘voltage drops’ are in the opposite direction of EMF and their sum is equal to EMF according to Kirchhoff’s second law.

6.6.5 Operating principles of electrical machines

The operation of electrical machines is explained by four general principles:

1. The electromagnetic action principle: An electrical current creates a surrounding magnetic field that is strengthened by passing through an iron core.
2. The motor action principle: An electrical current in a magnetic field (produced by some other source or electrical current) experiences a force perpendicular to both the direction of the current and the direction of the magnetic field, and reverses if either of these reverse in direction. The force is proportional to the current and to the strength of the magnetic field.
3. The generator action principle: A voltage (electromotive force) is induced in a conductor moved in a magnetic field. The induced voltage will cause a electrical current to flow. Note that the voltage is opposite to the electrical current causing a force in the direction of motion by principle 2. Generator action will only produce AC voltage, which must be modified if DC voltage is required.
4. The transformer action principle: A changing magnetic field induces a voltage. Only a change in the magnetic field induces voltage; if the magnetic field remains constant for any length of time, then no voltage will be induced (i.e., voltage = 0).

6.6.6 Rotating electromagnetic system elements

Rotating electromagnetic systems have two modes of operation: a motor mode and a generator (AC) mode.

Any given rotating electromagnetic machine may be designed to operate as a motor and/or a generator. Some of these machines, without any change in configuration, may operate as a motor and generator (but, not at the exact same time; e.g., induction motor/generator).

NOTE: *Every rotating electromechanical/ electrical machine is capable of working as a generator as well as a motor.*

There are three principal types of rotating electromagnetic system:

1. **DC motors**. - use DC current/power with electromagnetic induction to produce mechanical power.
2. **AC motors** - use AC current/power with electromagnetic induction to produce mechanical power.
3. **AC generators** (which may have their AC output converted to DC, and therein, may be referred to as **DC generators**) - AC generators use mechanical power with electromagnetic induction to produce AC current/power, and DC generators are AC generators with additional equipment to convert the AC current/power into DC current/power.

NOTE: *All rotating generators produce AC internally, and must have additional components to convert the AC into DC.*

All rotating electromagnetic systems have two categories of movement (i.e., two mechanical elements to which components are attached): the strator (stationary element) and the rotor (rotating element). They also have two categories of electrical elements: the armature (power producing component) and the field (the magnetic field component). A rotating electrical machine consists of a field and an armature where rotation occurs with respect to each other. The armature is the part of the machine in which the energy "conversion" takes place.

The mechanical elements of rotating electromagnetic systems are:

1. **The stator** - all of the non-rotating electrical parts of a machine (motor or generator). The stator is the outer shell of the motor or generator that remains stationary during operation.
   - Strator electrical element variations include:
   - **Strator armature winding/coil** - generated current to load.
   - **Strator field windings** (forming an electromagnetic electro motor) - AC or DC supplied.
   - **Strator permanent magnets** (strator-PM motor) - magnets mounted to strator.
   - Strator-fed commutator
2. **The rotor** - all of the rotating electrical parts of the machine (motor or generator). The rotor is the central spinning core of the motor or generator.
   - Rotor electrical element variations include:
   - **Rotor armature winding/coil** - generated current to load.
   - **Rotor field windings** (forming an electromagnetic electro motor) - AC or DC supplied.
   - **Rotor winding as cage** (rotor cage) - windings are shorted.
   - **Rotor permanent magnets** (rotor-PM motor) - magnets mounted to rotor.
   - **Rotor-fed commutator**
   - **Slip ring attachment**
3. **The [stator] air gap** - the gap between the stator and the rotor (the air gap separating the inner stator and outer rotor surfaces. A gap must exist for the rotor to rotate.

The electrical elements of rotating electromagnetic systems are:

1. **The armature** - the power-producing component of the machine. It is the main current-carrying winding/coil in which the electromotive force or counter-emf of rotation is induced. The armature has two functions: 1) to carry current crossing the field, thus creating shaft torque in a rotating machine or force in a linear machine; and 2) to generate an electromotive force (EMF).
   - The armature can be on either the rotor or the stator.
   - In other words, the armature winding/coil is that which generates or has an alternating voltage applied to it. The current in the armature winding/coil is known as the armature current. The location of the winding depends upon the type of machine -- it can either be part of the stator (strator coil) or the rotor (rotor coil) as long as voltage is induced. In the armature, an electromotive force is created by the relative motion of the armature and the field. When the machine acts in the motor mode, this EMF opposes the armature current, and the armature converts electrical power to mechanical power in the form of torque (unless the machine is stalled), and transfers it to the load via the shaft. When the machine acts in the generator mode, the armature EMF drives the armature current, and shaft mechanical power is converted to electrical power and transferred to the load.
   - In a generator, the windings from which current is generated are called armature windings. All other windings therein are field windings.
   - The armature always carries current; hence, it is

always a conductor or a conductive coil.

2. **The field -** the magnetic field (i.e., magnetic flux) component of the machine. It is the part that generates the direct magnetic field. The field can be on either the rotor or the stator and can be either an electromagnet (field coil) or a permanent magnet. A field coil is an electromagnet used to generate a magnetic field. It consists of a coil of wire (winding) through which a current flows (field winding or field coil).
   - The current in the field does not alternate. The field may either be stationary (magnets attached to the strator), or the field may rotate (magnets attached to the rotor). A rotating magnetic field is one whose north and south poles move inside the stator, just as though a bar magnet, or magnets, were being spun inside the machine.
   - The speed at which the magnetic field rotates is called the synchronous speed and is described by the following equation: S = (f x P) / 120 where S = rotational speed in revolutions per minute f = frequency of voltage supplied (Hz) P = number of magnetic poles in the rotating magnetic field.
   - The path of the magnetic field is determined by the presence of magnetic "poles", which are located [at equal angles] around the rotor/strator. At each "pole", magnetic field lines pass from stator to rotor or vice versa. The stator (and rotor) are sub-classified by the number of poles they have.

Rotating a conductor through a magnetic field generates an electromotive force (EMF, voltage). The same effect can be accomplished by holding the conductor stationary and rotating the field. Either the conductor or the field must remain stationary, or move at a slower rate. That is to say that relative motion must always exist between the armature and the field in order to generate an EMF in the armature. In either case, "lines of force" are being "cut" by the conductor, generating an EMF in the conductor. If the speed of an armature rotor and strator field were the same (i.e., both rotating at the same speed), there would be no induced EMF, no lines of force would be "cut", and no field built around the conductor; hence, no generator or motor action under those conditions.

> **CLARIFICATION:** *In an electromagnet, magnetism is generated by electrical current. Magnetism is present only while electrical current is flowing. An electromagnet generates heat, but the heat does not significantly affect the magnetism. The more the electrical current and winding turns, the more the magnetism. From a structural and physical standpoint there is no difference between an AC and a DC electro-magnetic coil. They both are made by wrapping of wires around a core. For DC electro-magnets the core is usually made of iron or steel. For AC electro-magnetic coils the "core" could be air. In a permanent magnet, magnetism is retained after a material is magnetized (by an electrical current or by a rapidly moving magnet). Note here that high heating stresses will cause permanent magnets to become [irreversibly] demagnetized.*

The loop ends of the armature can be closed (i.e., "short circuited"), thereby inducing a larger current in the loop, and creating stronger magnetic fields around the conductors.

> **NOTE:** *Coils of wire formed into 'windings' can be classified into two groups: armature windings and field windings. The armature winding is the main current-carrying winding in which the electromotive force (emf, generator) or counter-emf of rotation (motor) is induced. In other words, the armature winding is the winding (conductive coil) to which electrical current is supplied in the case of a motor (to induce rotation), or from which electrical current is extracted/generated in the case of a generator. The current in the armature winding is known as the armature current. The field winding produces the magnetic field in the machine (unless there is a permanent magnet, which produces the field). The current in the field winding is known as the field or exciting current. The location of the winding depends upon the type of machine.*

There are three primary ways of creating a changing/fluctuating (in flux) magnetic field:

1. A constant-magnitude magnetic field pattern is moved repeatedly in space past a stationary conductive path (e.g., a synchronous generator whose magnetized rotor poles move repeatedly past its stator windings/strator coils).
2. A path for an electromagnetic force (EMF) in space (a coil of wire, as in, rotor windings/coils; or other conductive material) is moved repeatedly past a constant magnetic field fixed in space (e.g., AC generator with a commutated armature to produce DC current).
3. A magnetic field that varies in both time (electromagnetic fluctuation) and in space (physical fluctuation) moves past a stationary conductive path (a strator winding/strator coil). Here, currents are induced in the rotor and create a changing magnetic field that sweeps repeatedly past the stationary stator windings. Depends on an external voltage source to produce the electromagnetic fluctuations.

The source of the magnetic field for #1 and #2 can be either one or more permanent magnets or externally supplied currents in coils of wire. Permanent magnet generators exist in contrast to generators with field windings, which have "field circuitry", and require an external power sources to pass electric current through the field windings (to create an magnetic field via electromagnetism).

6.6.7 DC voltage [power] Generation

There are a number of ways of producing direct current voltage. A potential difference (voltage) between which direct current flows can be generated in the following ways:

1. **Rectification [of alternating current]** - The task/process of converting (i.e., rectifying) alternating current into DC current is known as rectification. A **rectifier** (literally, "to make straight") is an electrical device that converts alternating current (AC), which by definition periodically reverses direction, to direct current (DC), which flows in only one direction. A rectifier contains electronic elements (usually) or electromechanical elements (historically) that allow current to flow only in one direction - alternating current is supplied and direct current is output.
    - *Rectifier AC -> DC*
    - Take note that all 'generators' produce AC as their first internal output, which may then be rectified by another (external or directly incorporated) device (e.g., commutator) to produce DC. A 'DC generator' is an 'generator' (AC), with a rectifier attached to it (or incorporated into it).
    - **Commutation** - Commutation is a type of rectification used in rotating electromagnetic generators. Commutation uses the positioning of conductive elements (contact bars or metallic brushes) connected to the rotor to convert the armature's AC input to a DC output by changing direction at the same time the incoming armature AC changes direction. A commutation device/mechanism is known as a commutator. The commutator serves to "rectify" the induced AC voltage in the armature. Commutation changes the direction of current so that the system's output always experiences "somewhat continuous" EMF (i.e., does not experience an alternating current voltage). In general, commutation produces a type of pulsating electromotive force (pulsating DC) that can be "smoothed out" by additional electrical techniques to give a sufficient imitation of direct current (essentially, DC). In other words, when the shaft of an AC motor is mechanically coupled to a commutator (creating a DC generator), then DC is output.
2. A **homopolar generator** (a.k.a., **unipolar dynamo** or **Faraday disk**) - a unique type/configuration of rotating electromagnetic generator that produces a direct current without the need for a rectification (e.g., commutator), using a copper disc rotating within a magnetic field. This setup produces a low DC voltage, and high amount of current.
3. **Photovoltaics (PV)** - A photovoltaic cell is simply a semiconductor device made of P-type (positive charge) and N-type (negative charge) materials. The boundary between P and N acts as a diode allowing electric charges to move from N to P, but not from P to N. When light with sufficient energy makes contact with the N-type material, electric charges move toward the P-type material creating a voltage difference which results in current generation. The current generated is direct current.
4. **Osmotic power** - Osmotic power can be used to produce DC through the use of a semi-permeable membrane that separates two fluids with different solutions. Ions travel through the membrane until the ionic concentrations in the two fluids reach equilibrium. When the ions pass through the membrane, charges are transferred to an electrode and produce DC voltage.
5. **Batteries** - a chemical reaction inside of the battery produces DC voltage at its terminals.
6. **Fuel cells** - A fuel cell is a device that converts the chemical energy from a fuel into DC voltage through a chemical reaction of positively charged hydrogen ions with oxygen or another oxidizing agent. Fuel cells are different from batteries in that they require a continuous source of fuel and oxygen (or air) to sustain the chemical reaction, whereas in a battery the chemicals present in the battery react with each other to generate an electromotive force (emf). Fuel cells can produce electricity continuously for as long as these inputs are supplied.
7. **Capacitors** - once charged, capacitors give a regulated DC supply. In an energized circuit, however, capacitors will "block" DC (Note: there are a variety of uses for capacitors in circuits).

NOTE: *Motor-generator set (M-G set), which combine a motor's mechanical output connected to the rotor of a generator may be used to create DC voltage and/or modify existing DC voltage. Such devices have the*

- (1) To convert from AC to DC. An AC powered motor connected to a DC generator.
- (2) To modify DC voltage - DC at a fixed voltage to DC at a different voltage.
- (3) To create or balance a 3-wire DC system.

Also, generally, the term **dynamo** *(or dynamotor) refers to a DC only motor-generator set (i.e., DC input and DC output).*

6.6.8 AC voltage [power] Generation

There are three ways of producing alternating current voltage: (1) conversion of DC current via an inverter (inversion-type); (2) an AC generator (alternator-type or induction-type); and/or (3) a motor-generator set, which is an AC generator with a motor connected to its rotor/shaft.

1. A **power inverter** (or **inverter**) is a device that converts direct current (DC) into alternating current (AC). The input voltage, output voltage and frequency, and overall power handling depend on the design of the specific device or circuitry. The inverter does not technically "produce" power; the power is provided by the DC source, and flows through to its AC output. One means of changing from direct to alternating current is to use a motor-generator set (M-G set) as an inverter. The converted AC can be output at any required voltage and frequency with the use of appropriate transformers, switching, and control circuits.
    - *Inverter DC -> AC*

2. An **AC generator** (alternator and induction generator) converts mechanical energy into AC electrical energy (as its output) based on the principle of electromagnetic induction. All "generators" produce AC internally. It requires a moving magnetic field and a conductor, which together form an electromotive force (EMF) in the conductor. Take note that a simple generator without a commutator will produce an electric current that alternates in direction as the armature revolves -- with a commutator it will produce DC current. Alternator technology may be classified by method of excitation, number of phases, the type of rotation, and their application.

    AC can be produced using a device called an alternator. This device is a special type of electrical generator designed to produce alternating current. Typically, a rotating magnet, called the rotor turns within a stationary set of conductors wound in coils on an iron core, called the stator. The rotating magnetic field induces an AC voltage in the stator windings. Since the currents in the stator windings vary in step with the position of the rotor, an alternator is a synchronous generator. A loop of wire is spun inside of a magnetic field, which induces a current along the wire. The rotation of the wire can come from any number of means: a wind turbine, a steam turbine, flowing water, and so on. Because the wire spins and enters a different magnetic polarity periodically, the voltage and current alternates on the wire.

    The speed at which the rotor spins in combination with the number of generator poles (i.e., magnetic "poles" in the generator) determines the frequency of the alternating current produced by the generator. All generators on a single synchronous system, for example the national grid, rotate at sub-multiples of the same speed and so generate electric current at the same frequency. If the load on the system increases, the generators will require more torque to spin at that speed and, in a typical power station, more steam must be supplied to the turbines driving them. Thus the steam used and the fuel expended are directly dependent on the quantity of electrical energy supplied. An exception exists for generators incorporating power electronics such as gearless wind turbines or linked to a grid through an asynchronous tie such as a HVDC link — these can operate at frequencies independent of the power system frequency. Depending on how the poles are fed, alternating current generators can produce a variable number of phases of power. A higher number of phases leads to more efficient power system operation but also increases the infrastructure requirements of the system

    AC can come in a number of forms, as long as the voltage and current are alternating. If we hook up an oscilloscope to a circuit with AC and plot its voltage over time, we might see a number of different waveforms. The most common type of AC is the sine wave. The AC in most homes and offices have an oscillating voltage that produces a sine wave. Other common forms of AC include the square wave and the triangle wave. Triangle waves are found in sound synthesis and are useful for testing linear electronics like amplifiers.

3. An **AC motor-generator set** (a.k.a., **M-G set; MG; engine-generator; gen-set; generator**) is a combination motor and generator system for converting electrical current from one form to another (or converting voltage and or frequency of the same current or between two different currents). Such devices may also be used to isolate electrical loads from an electrical power supply line. Herein, an electrically powered motor (either DC or AC powered) is mechanically connected to an AC

generator. A motor-generator set involves a motor operating from an available electric power supply to drive a generator, which delivers (outputs) the current and voltage wanted (external power runs motor > motor powers generator > generator outputs desired power). The motor is mechanically coupled to an appropriate generator, creating the desired conversion. In AC applications, an M-G set has four possible functions:

A. To convert from DC to AC (as an inverter). A DC powered motor connected to an AC generator.
B. To modify frequency - AC at one frequency to AC at another harmonically-related frequency.
C. To modify voltage - AC at a fixed voltage to AC of a variable voltage.
D. To modify phase - AC single-phase to AC three-phase.

The mechanical torque required to power a generator may come from the following sources:

1. Animate power (animal movement generates shaft power; animal-generator)
2. Turbine power (fluid movement, gaseous expansion or combustion spins a propeller-like device and generates shaft power; turbine-generator)
3. Motor power (a motor or engine generates shaft power; motor-generator)

Voltage and frequency regulation in a generator is maintained by controlling:

1. Generator excitation
2. The speed of the prime mover
3. Shaft speed through a gearbox
4. Electronically

Aside from the internal configuration of a generator, the frequency expressed by all generators depends upon the rotational speed of the generator's shaft (rotor) and the load. A faster rotation of the shaft will generate a higher frequency. A higher load will slow the rotor, possibly to the point where it reduces the frequency. At no-load, the mechanical system is rotating at the 'no-load speed', and results in the generation of voltages at 'no load frequency'. When a generator is loaded, power is drawn from the mechanical system and the generator experiences a torque that opposes the direction of motion of the mechanical system. As a result, the mechanical system of the generator tends to slow down.

6.6.9 AC voltage generation: phase

CLARIFICATION: *Phase, like frequency, is a concept restricted to AC voltage generation (and does not apply to DC voltage generation).*

AC voltage may be sub-classified by phase, wherein there are three principal types of AC voltage: single phase, two phase, and polyphase (e.g., three-phase). Systems with more than two phases are generally termed polyphase. Polyphase systems have three or more energized electrical conductors (three or more phases) carrying alternating currents with a definite time offset between the voltage waves in each conductor. Polyphase systems are particularly useful for transmitting power to electric motors. Once polyphase power is available, it may be converted to any desired number of phases with a suitable arrangement of transformers. Conversion between polyphase systems of different phase numbers is always possible. Polyphase systems are qualitatively different from single phase systems. Note here that the order of voltage waveform sequences in a polyphase system is called *phase rotation* or *phase sequence*.

NOTE: Phase converters *convert between different AC phases.*

There are several basic types of AC voltage generation:

- **Single phase AC voltage** are defined by having an AC source with only one voltage waveform. In other words, there may be more than one voltage, but all voltage waveforms are in phase, or in step, with each other. Here, when more than one phase is present, the currents in each conductor reach their peak instantaneous values sequentially, not simultaneously. Note that a single phase supply connected to an alternating current electric motor does not produce a revolving magnetic field; single-phase motors need additional circuits for starting, and such motors are uncommon above 10 kW in rating.
  - Single-ended single-phase system: 1 phase, 2 wire - one of the wires is for the power, and one wire is for neutral.
  - Split-phase (single-phase three-wire): 2 phase, 3 wire - two of the wires are for phases (phase A, phase B), and one wire is for neutral.
- **Two phase AC voltage** by having voltage phases differing by one-quarter of a cycle, 90°. Usually circuits used four wires, two for each phase.
  - 2 phase, 4 wire - two separate pairs of current carrying conductor, and no neutral.
  - 2 phase, 3 wire - two wires carry two separate phases, and the common conductor (wire) carries the vector sum of the phase currents, which requires a larger conductor. No neutral.
  - Note: Two-phase power can be derived from a three-phase source using two transformers in a Scott connection.
- **Three phase AC voltage (polyphase)** are defined by having three or more energized electrical conductors carrying alternating currents with a

definite time offset between the voltage waves in each conductor. All 3-phase generators (or motors) use a rotating magnetic field. A polyphase power system uses multiple voltage sources at different phase angles from each other (many "phases" of voltage waveforms at work). A polyphase power system can deliver more power at less voltage with smaller-gage conductors than single- or split-phase systems. The phase-shifted voltage sources necessary for a polyphase power system are created in alternators with multiple sets of wire windings. These winding sets are spaced around the circumference of the rotor's rotation at the desired angle(s). A major advantage of three phase power transmission (using three conductors, as opposed to a single phase power transmission, which uses two conductors), is that, since the remaining conductors act as the return path for any single conductor, the power transmitted by a balanced three phase system is three times that of a single phase transmission but only one extra conductor is used.

- 3 phase, 4 wire - three wires for the power and one for neutral.
- 3 phase, 4 wire Delta - three for the power and one for neutral.

6.6.10 AC voltage generation: synchronous and asynchronous speeds

AC machines (generators and motors) can be divided into two main categories: synchronous speed machines and asynchronous (induction) speed machines. In concern to generators, synchronous speed AC voltage generators are commonly referred as alternators. And, asynchronous speed AC voltage generators are also known as 'induction generators'. Regardless of naming, both synchronous and asynchronous devices (motors and generators) use electromagnetic induction as their primary operational effect.

The 'synchronous speed' is that which causes the generator to produce the grid frequency exactly. If the grid frequency is constant, so is the 'synchronous speed'. Asynchronous means that the machine cannot produce torque (motor) or power (generator) when turning at the synchronous speed. To emphasize, an asynchronous machine cannot operate at the synchronous speed. In an asynchronous machine, when the rotor rotates at synchronous speed, no interaction takes place between magnetic field and the rotor because they are moving together (creating the condition of 'zero slip'), and thus, no torque or power will be induced. This difference between the actual speed of the rotor and the synchronous speed is called the 'slip'.

In an asynchronous machine, when the rotor rotates faster than synchronous speed, it inputs electrical power into the power network as 'positive slip', and when it operates below synchronous speed it acts as a load and pulls power from the power network as 'negative slip'. Conversely, a synchronous generator operates at exactly the same frequency as the [power] network to which it connects.

> **EXAMPLE OF SYNCHRONOUS SPEED:** *For a typical four-pole induction machine (two pairs of poles on stator) operating on a 60 Hz electrical grid, the synchronous speed is 1800 rotations per minute (rpm). Hence, the machine must rotate faster than 1800rpm to begin generating electrical power.*

This difference of rotor speed from magnetic field speed in both motoring and generation is referred to as positive and negative slip, respectively.

- Slip = $(f_0 - f_r) / f_0$
- where:
  - f_0 = frequency of the electrical grid (synchronous speed in revolutions per minute)
  - f_r = frequency of the rotor (rotor speed)
  - (at start-up slip = 1, at synchronous-speed slip = 0)

6.6.11 Synchronous generators (alternators)

A.k.a., Syn-chronous (same-time) generators of electrical power.

Synchronous generators (SG; alternators) are called "synchronous" because the waveform of the generated voltage is synchronized with the rotation of the generator -- there is no phase shift, the speed of the rotor is called synchronous speed (constant speed) -- the rotor and magnetic field rotate at the same speed. Each peak of the sinusoidal waveform corresponds to a physical position of the rotor. Synchronous motors and generators are nearly identical.

For a synchronous generator, frequency is determined by the rotational speed of the generator's shaft -- faster rotation of the shaft generates a higher frequency. In other words, if the electrical output frequency of the generator is synchronised to its shaft/rotor speed, then it is a synchronous generator system. The synchronous generator's rotational speed is locked to its stator frequency. However, the electrical output frequency is not necessarily synchronised to the grid frequency. External controls may be necessary to achieve the correct grid frequency.

In the majority of designs, the 'rotor' contains the magnet (rotating field), and the 'stator' is the stationary armature (armature windings) that is electrically connected to a load. The magnetic field source (magnetic flux) may be supplied by either permanent magnets (in the rotor) or an excitation current fed into field windings (in the rotor):

1. **Permanent-magnet synchronous generator (PMSG)** has permanent magnets (permanent magnet excitation). Also known as **permanent magnet alternator (PMA)**. PMSG's are simpler and do not consume/require power to generate the field flux.
2. **Wound field synchronous generator (FESG)** has direct current flowing in wound field windings (wound field excitation) to create an electromagnet. If the field winding is directly connected across the armature output terminals to obtain its power, it is called 'shunt excitation'. If the field current is controlled separately from the armature voltage, it is called 'separately excited'. FESG's allow for greater control, but require power (DC voltage) to generate the field flux (field current). The torque and output power of a wound field generator can be controlled by adjusting the field current (electromagnets) of the generator.

The direct current in the rotor field winding is fed through a slip-ring assembly or provided by a brushless exciter on the same shaft. Thus, electrical power is generated by moving rotor and its attached permanent magnet or electromagnetic within a stationary casing that contains armature windings around the outside, which are electrically connected to a load. When an electromagnet is used, it draws its excitation from a power source external to, or independent of, the load or transmission network it is supplying.

NOTE: *Slip rings act as load connectors (i.e., they connect the armature winding(s) to an external load).*

When voltage is applied to the armature windings, then electromagnetic induction causes the rotor (field) magnet to spin/rotate (creating a synchronous motor). Conversely, when the rotor with the permanent magnets are spun through mechanical power applied to rotate the shaft, then electromagnetism induces an a voltage as an alternating current in the armature windings (synchronous generator or alternator).

The expression relating the rotational speed in revolutions per minute, the number of magnetic poles in the machine, and the electrical frequency in hertz can be expressed as:

- rpm = 120 x frequency/poles

NOTE: *There are fixed-speed synchronous generators and there are variable-speed synchronous generators.*

6.6.12 Induction generators (asynchronous generators)

A.k.a., Asyn-chronous (alternating-time) generators of electrical power.

Induction generators (IG; asynchronous generator, AG) are essentially the same machine as an asynchronous or induction motor - an induction generator is mechanically and electrically similar to an induction motor. The principle of operation of the induction motor (and hence, generator) is based on generating a rotating, constant magnetic field. This rotating magnetic field interacts with a set of short circuited conductors arranged on the rotor. In other words, an induction generator is a type of electrical (AC voltage) generator that is mechanically and electrically similar to an induction motor. Hence, a regular asynchronous motor can usually be used as an asynchronous/induction generator, without any internal modifications.

Induction generators produce electrical power when their rotor shaft is rotated faster than the synchronous speed (frequency) of the equivalent induction motor. The rotating magnetic field induces currents in a set of copper loops in the rotor, and magnetic forces on these current loops exert a torque on the rotor and cause it to rotate (as a motor). When it is forced (from an outside energy source) to rotate past the synchronous speed, then it becomes a generator. In other words, they produce AC voltage when their rotor runs (moved by an outside mover such as wind) above the synchronous speed of the supplied voltage frequency. This requires an external torque applied to the rotor to turn it faster than the synchronous speed. If the generator's rotational speed is greater than the synchronous speed (the speed at which the magnetic field rotates), power is produced; if the speed drops below the synchronous speed, the generator becomes a load.

Summarily, induction machines have two operational modes:

1. Motor operation - When the current is connected, the machine's strator windings, the rotor will start turning like a motor at a speed which is just slightly below the synchronous speed of the rotating magnetic field from the stator. A phased induction motor works on the principle of electromagnetic induction, where the relative motion between the flux and the rotor, caused by the rotating magnetic field induces a current in the rotor, forcing it to rotate in the same direction. In an induction motor, the rotor rotates because of "slip" (i.e. relative velocity between a rotating magnetic field and the rotor). In order to maintain relative EMF, there must be a "slip" in the induction motor, or else the motor will stop. The rotor of the induction motor does not rotate as fast as the rotating AC [electro] magnetic field. In other words, the rotor goes slower than the rotating magnetic field in order to have relative motion. Within a 3 phase induction motor, motion is achieved by orientating the three electromagnetic coils (magnetizing flux) 120

physical degrees apart in space, and imposing 3 phase voltages on the windings also separated in time by 120 electrical degrees.

2. Generator operation - When the rotor moves faster than the rotating magnetic field from the strator, the strator induces a strong current in the rotor. The faster the rotor rotates (turning force, moment, or torque), the more power will be transferred as an electromagnetic force to the stator, and in turn converted to electric current (fed into an electrical grid).

> **CLARIFICATION:** *The strator contains electromagnets, and the rotor may simply be conductive and/or may contain permanent magnets. When current is supplied to the strator creating electromagnets, then the rotor will spin as a motor; but, when the rotor spins above synchronous speed due to a sufficient supply of outside power, then the strator windings will have current induced in them and produce [asynchronous] current on the circuit.*

Induction generators may be classified according to whether the rotor contains permanent or wound field windings, and whether the rotor's conductor is excited (i.e., energized, electrified):

1. When the rotor contains permanent magnets, then the system is called a permanent-magnet asynchronous generator (PMAG) has permanent magnets (permanent magnet excitation). PMAG's are simpler.
2. When there are no permanent magnets on the rotor, then the system is called a wound field asynchronous generator (FEAG) has direct current flowing in a wound field winding (wound field excitation) to create an electromagnet. FEAG's allow for greater control.
3. The term, 'doubly-fed induction machine' applies to a system where both the stator and rotor winding of a slip-ring machine are supplied with voltage (electrical power).

When the strator field windings are electrically excited, they behave like electromagnets, producing independent (per coil/winding) [electro]magnetic fluxes. The position of the [electro]magnetic fluxes around the strator keeps changing with time in a circular manner. Whereupon, net flux (resultant of all magnetic fluxes in the stator) develops a rotating magnetic field in the stator, which causes relative motion between the net flux (strator) and the rotor (current flows in the rotor winding) -- the rotor moves (i.e., rotates) as the magnetic flux in the strator rotates. The direction of rotation of the rotor is the same as that as the rotating magnetic field of the stator.

Induction generators (and motors) are not self-exciting; they require an external electrical supply to produce a rotating magnetic flux, and thus, induce current in the rotor. The electrical [supply] power required for this is called reactive current/power (i.e., they require reactive power for excitation). The induction generator depends on an external voltage source to produce a magnetic field (electromagnet) in the stator, which is to say that it consumes VARS (volt-ampere, reactive) in order to produce power (watts). In other words, an induction machine requires externally supplied armature current to start, and cannot start on its own as a generator. A source of excitation current (reactive power) is required to maintain the [electro]magnetic field (i.e., magnetizing flux) that induces current in the rotor. The excitation current supply can originate from: 1) the electrical grid; 2) from the generator itself (once it starts producing power); or 3) from a capacitor bank. If an induction generator is meant to supply a standalone load, a capacitor bank needs to be connected to supply reactive power. In other words, asynchronous machines are capable of self-excitation when, in order to supply the magnetizing current, capacitors are connected parallel to the machine terminals. Once the rotor reaches a speed above the armature currents supplied frequency, it will begin producing current.

Induction generators do not need to be synchronized with the grid before being connected. The generator is simply connected at dead standstill and grid power is used to operate the generator as a motor (at first), bringing it up to synchronous speed, whereupon it becomes a generator. Power is transmitted to the grid as long as the system turns faster than synchronous speed. Below synchronous speed, the generator acts as a motor and will consume power.

An asynchronous generator with an electronic controller can be allowed to vary with the speed of the energy source (e.g., wind). The output frequency and volts are regulated by the power system and are independent of input mechanical speed variations.

> **NOTE:** *Unlike synchronous generators, induction generators are load-dependent and cannot be used alone for grid frequency control. Wind turbine induction generators cannot support the electrical grid's system voltage during faults, unlike synchronous generators.*

There are two kinds of induction generators used in wind turbines:

1. Cage rotor induction generator/machine (a.k.a., squirrel cage rotor, SCIM) - has rotor windings (a cage winding), which are shorted (connected to themselves) and strator windings, which are connect to the grid or another power source. The rotating magnetic field in the strator induces a very strong current in the rotor bars, which offer very little resistance to the current, since they are short circuited by the end rings. The rotor then develops its own magnetic poles, which in turn become

dragged along by the electromagnetic force from the rotating magnetic field from the stator.

2. Wound rotor induction generator/machine (WRIM) - has rotor windings (connected to a load or power converter) and strator windings connected to the grid. Slip rings (and brushes) are used as parts of the rotor current. If the rotor coil windings were short circuited, then this machine would be similar to the cage induction machine; however, the rotor conductor cross section geometry is still different than that of a cage rotor induction machine. Induction machines with a wound rotor allow access to the rotor winding via slip rings and brushes. In other words, WRIMs require “slip rings” and brushes to supply electrical power (and resistance), whereas other induction machine configurations do not.

> **MAINTENANCE:** *Brush wear comes from two basic causes: mechanical friction and electrical wear. Mechanical friction is caused by the rubbing of the brushes on the commutator or slip ring. Electrical wear is caused by the arcing and sparking of the brush as it moves over the commutator. Mechanical friction increases with brush pressure; electrical wear decreases with brush pressure. For any given brush installation, there is an optimum amount of brush pressure. If the pressure is decreased below this amount, the total wear increases because the electrical wear increases. If the pressure is increased above the optimum amount, the total wear again increases because mechanical friction increases.*

If an induction generator is supplying a standalone load, the output frequency will be slightly lower (by 2 or 3%) that calculated from the formula:

- f = N * P / 120.
- where, N is speed of the rotor in rpm and P is number of poles.

6.7 Voltage conversion and inversion

The process of changing AC voltage into DC voltage is called conversion (actually, this is an imprecise term because “conversion” also refers to changing one DC voltage to another, and other things as well, but it will do for our purposes). Devices that perform this process are called converters, but are also sometimes called [power] ‘adapters’, and if being used for charging batteries, they are often just called [power] ‘chargers’. Changing DC into AC is the opposite process and is called inversion. A device that does this is, of course, called an ‘inverter’.

A transformer is an electromagnetic device that changes (or “transforms”) AC current at one voltage (in one circuit) to AC current at another voltage (in another circuit). In the simplest case, most transformers consist of a metal rectangular core, around two sides of which two separate wires are wound each connected to a separate circuit. The rectangular core is generally iron (ferromagnetic), and hence, nearly all the flux from the first circuit will be transferred to the secondary circuit’s windings (inducing current in the secondary circuit). That which makes the transformation is the difference in the number of coil/winding turns on both sides.

> **NOTE:** *When a transformer is present, then two electrical subsystems are created, because electrical current on one side of a transformer does not flow into the circuit on the other side. There is a physical coupling between the two subsystems, but no direct electrical connection. The transformer becomes an interface between the two subsystems.*

For a transformer to work, the current in one coil has to somehow make current flow in the other coil (and the circuit it’s connected to). A DC current in one coil will make a magnetic field on the other coil, but a magnetic field by itself won’t drive any electrons around (electromotive force is not produced). A *changing* magnetic field (i.e., time-varying magnetic field), however, does create an electric force, which accelerates the electrons in the other coil into carrying a current. This process is described by Faraday’s law of induction. AC current produces a changing field, because the current which makes the field is changing.

Transformers work via induction of electrical forces by changes in magnetic fields. Both AC and DC generate a magnetic field. However, because DC currents produce a constant magnetic field, their passage through a transformer will not generate an electromotive force in the secondary circuit; however, it will still be “consuming” energy. In AC, when the current changes direction, so does the field, which causes an EMF in the secondary circuit and moves charges therein.

A DC to DC (DC/DC) converter can be described as the DC equivalent of an AC transformer. It changes the ratio between the input and output voltages and currents by introducing ‘power electronics’ that, with the help of passive components, transmit the power through the converter. These solid state devices, which are products of the semiconductor revolution, make it possible to transform DC power to different voltages. The advantages of using DC/DC converters are many: To regulate the output voltage, to build subsystems supplied by the same bus and to reduce transmission losses.

6.8 Electric power transmission & distribution (transportation)

> **REMEMBER:** *Every transfer between energy carrier (“conversion”) represents a loss of energy.*

Electrical power transmission refers to the movement of electrical power from one location to another. Electrical power transmission refers to the bulk movement of electrical energy a significant spatial distance from a generating site, such as a power plant, to a point where it is distributed for end-use/service. Thus, electrical power distribution refers to the distribution of electrical power a relatively short distance (in comparison to transmission) from a source location its end-use/service. Electrical power generated proximate (i.e., near) to end-use does not involve long-distance transfer (transmission), but will still require short-distance transfer (distribution).

Presently, there are two primary forms of electric power transmission & distribution, categorized by their medium of transmission.

1. **Wired** - the transmission (and subsequent distribution) of electricity with the use of wires or other conductive guiding structures that form an electrical power network ("the grid").
2. **Wireless** - the transmission of electricity without wires or other guiding structures.
3. **Storage** - the storage of energy in a carrier, which may be transported, and then easily transferred through to electrical power.

6.8.1 Wired electric power transmission and distribution

Electric power transmission is the bulk movement of electrical energy from a generating site, such as a power plant, to a point where it is distributed for end-use/service. The interconnected [conductive] lines which facilitate this movement are known as a [electrical power] 'transmission network' (long-distance transmission). This is distinct from the local wiring between high-voltage substations and customers, which is typically referred to as [electric power] 'distribution network' (short-distance transmission). The combined transmission and distribution network has several names, including but not limited to: "power grid"; "the grid"; and the "National Grid".

> **CLARIFICATION:** *Electrical power transmission is the transfer of electric energy as electrical power over an interconnected group of conductive lines and associated equipment between points of supply and points at which it is transformed for delivery to end use (or other electric systems). Transmission is considered to end when the energy is transferred for distribution to end use (e.g., at a substation).*

The best way to transmit power (both AC and DC voltage) efficiently is to send it at very high voltage and very low current: high voltage AC (HVAC) or high voltage DC (HVDC). Current is affected by line resistance (impedance), and so, it is necessary to send very little amperage to reduce power loss from heat. Once the voltage is increased to a very high voltage, then there is no inherent advantage to its being AC or DC. The generated electric power is often stepped up (at a step-up transmission station) to a higher voltage, whereupon it connects to an electric power transmission network. On arrival at a step-down substation, the power will be stepped down from a transmission level voltage to a distribution level voltage for distribution to end use -- as it exits the substation, it enters the distribution wiring. Finally, upon arrival at the service location, the power is often stepped down again [in voltage] from the distribution voltage to the required service voltage(s).

> **NOTE:** *All "modern countries" are criss-crossed with high-voltage transmission lines, which transport electrical power from generators at power plants to substations and ultimately consumers. This is partly because most electrical power generation systems are located away from population densities, and also, partly because are so spread out (i.e., 'population sprawl').*

Hence, most wired electric power transmission and distribution systems ("grids") consist of the following components, listed in order from generation to end-use:

1. Power station - power supply (generation).
2. Step-up transmission substation - steps up the voltage for transmission.
3. Transmission network - conductive lines for transmission across a significant spatial distance (long-distance transmission).
4. Step-down transmission substation - steps down the voltage for distribution.
5. Distribution network - conductive wires for distribution to end use/service (short-distance transmission).
6. End-use voltage transformers (AC)/converters (DC) - step down the distribution voltage to the voltage required for individual electrical devices.
7. Load - the end-use/service.

High-voltage transmission lines transport power over long distances much more efficiently than lower-voltage distribution lines for two main reasons. First, high-voltage transmission lines take advantage of the power equation: power is equal to the voltage times current ($P = VI$). Therefore, increasing the voltage allows a decrease in current for the same amount of power. Second, since transport losses are a function of the square of the current flowing in the conductors, increasing the voltage to lower the current significantly reduces transportation losses. Additionally, reducing the current allows for smaller transmission conductor sizes. As the length of any conventional transmission line increases, both the energy transfer capacity of the line and the efficiency of energy transfer decrease. The primary ways in which to overcome energy loss are to increase the transmission line voltage, and/or to increase wire diameter.

> **HISTORICAL NOTE:** *The reason AC is the primary power transmission and end-point access goes back to the late 1800s. Back then there were two competing power grids, one for*

AC and one for DC. Thomas Edison developed DC power, and Nikola Tesla developed AC. With the technology available at that time, Tesla's AC power could be transported long distances more efficiently and with lower cost. At the time, the invention of the electrical transformer meant that AC electrical power could be stepped up relatively easily to a higher voltage and transmitted more efficiently (at a higher voltage, but lower current). Efficient electrical transmission meant in turn that electricity could be generated at centralised power stations, where it benefited from economies of scale, and then be despatched relatively long distances to where it was needed. Hence, in the market, AC won out and became the only grid available. At the time that the primary power grids linking major cities on this planet were first built, starting around the 1880s, there was no convenient way to change the voltage of DC power, whereas the materialized presence of a "transformer" made AC voltage modification possible. Transformers do not work for DC (there are no DC transformers). Up until 1956, only AC power could be readily changed from one voltage to another (via transformers). the voltage level that came out of the DC generator was essentially all you had to work with and that voltage was usually too low to transmit power very far without substantial losses. Hence, AC is the dominant form of power today for the simple reason that when these grids were first being setup, the technology required to easily and cheaply manipulate DC voltages did not exist. Today, the universal standard power grid on earth is alternating current. With the technology available today, high voltage DC (HVDC) is the optimal way of transmitting large amounts of power great distances.

In power systems where generation is distant from the load, AC voltage (or DC) can be stepped-up (increased) at the generation or transmission point, and then, step-down (decrease) the voltage near the load. Conversely, generators can be designed produce higher voltages.

AC power has the advantage of being easy to transform between voltages and is able to be generated and utilised by brushless machinery. Stepping DC voltage up and down (for high voltage transmission lines) is a more complex issue than AC voltage stepping (using transformers). However, DC can be more economical to transmit over long distances at very high voltages (via HVDC). In voltage transmission, compared to the charge movement of DC, AC power is inefficient due to the energy radiated (i.e., lost) with the rapid reversals of the currents polarity. We often hear these reversals as the familiar 50 or 60 cycles per second (50 or 60 hertz) hum of the appliance. AC power is also prone to harmonic distortions, which occur when there is a disruption in the ideal AC sinusoidal power wave shape, which also happens to be a wave shape lethal at sufficient voltage if it crosses the human heart.

NOTE: *Transmission of any current over long distances requires lethally high voltages.*

Presently, AC suffers from a variety of problems created by line impedance (X of C and X of L), which does not affect DC. AC suffers from losses due to "skin effect", as well as dielectric losses, that typically limit the voltage it can be increased to, which is approximately 765,000Vrms (765kV). Dielectric losses are caused when dipoles in matter align with a changing local electric field. As the polar structures turn to follow the field, the movement causes local heating. AC requires special wire, and wiring techniques, to control these losses. DC, however, can be stepped up to much higher voltages. Presently, transmission voltages as high as 1,600,000VDC, are being used. This means that high voltage DC (HVDC) can be sent more efficiently, at present, than AC. There are other technical reasons why AC suffers certain limitations in certain applications like underground power transmission, or that AC line power must be synchronized with the local AC grid at both ends of the line, whereas DC power can bridge between two different synchronized AC grids that are not synchronized with each other.

NOTE: *DC power remains the only practical choice in digital electronic systems. In other words, they require direct current for the circuit to correctly complete. Hence, with an AC distribution grid, a technology is needed to convert the current to DC. This technology is commonly known as a 'power supply', and they may be built into, or an external attachment to, a DC electronic device. Conversely, a technology known as an "inverter" changes a direct current (DC) into an alternating current (AC).*

In general, the distribution lines of an electrical grid are passive systems (i.e., they are not actively managed by operators or computer programs). Also, since they are traditionally unidirectional in power flow (from high voltage to low voltage), they do not require much maintenance.

When a part of the network or grid connects to another part of the network or grid, the area where the two or more grids/networks connect is called a grid 'interconnection'. Grid interconnection variables include but are not limited to:

1. Frequency regulation (AC only)
2. Voltage regulation
3. Disconnection and reconnection protocols
4. Safe intentional islanding operation
5. Control of faults

6.8.2 High-voltage AC and DC Grids

High voltage AC (HVAC) and high voltage DC (HVDC) electrical grids have different requirements due inherent differences in the characteristics of their currents.

1. **HVAC Grids:** AC grids must maintain steady frequency and voltage levels to avoid damaging demand-side equipment. Hence, they must actively limit harmonics, which are distortions of the normally smooth sinusoidal variation of an AC grid's voltage. Harmonics contribute to system inefficiency: they decrease the efficiency of motors by their inability to contribute to motor torque; they result in the heating of motors; they cause unbalanced currents in power systems; and they can damage electronic and computer components.
   A. Frequency regulation: When large generators are connected to the grid, they will set the grid's frequency. Therein, small generator do not have to regulate their own frequency.
   B. Voltage regulation: Whereas frequency is a variable that is constant across the whole utility electric power system (and thus subject to control throughout the system by a few large generators) voltage varies from node to node throughout the system depending on the distribution of loads, generation, and power factor correcting capacitor banks.
2. **HVDC grids:** DC grids are concerned with maintaining steady voltage, and the notion of frequency and harmonics do not apply. There is no such thing as DC harmonics, as DC is defined as zero frequency (zero sequence harmonic).

6.8.3 The wired electrical/power grid

An electrical/power grid is an interconnected network for delivering electricity from its point(s) of generation to its point(s) of usage/demand. The term 'grid' usually refers to a network, and should not be taken to imply a particular physical layout or a scale. The word 'grid' may also be used to refer to an entire continent's electrical network, a regional transmission network, or it may be used to describe a sub-network, such as a local utility's transmission grid or distribution grid. Electricity grid systems connect multiple sources/generators and loads. In electrical grids, a power system network integrates transmission grids, distribution grids, distributed generators and loads that have connection points called buses.

The structure, or "topology", of a grid can vary depending on the requirements of the system, including reliability, and the load and generation parameters. The physical layout is often forced by what land is available and its geology. There are multiple types topologies, including:

1. **Radial network topology** - The simplest topology for a distribution or transmission grid is a radial structure. This is a tree shape where power from a large supply radiates out into progressively lower voltage lines until the destination homes and businesses are reached. Most transmission grids offer the reliability that more complex mesh networks provide. The expense of mesh topologies restrict their application to transmission and medium voltage distribution grids. Redundancy allows line failures to occur and power is simply rerouted while workmen repair the damaged and deactivated line. A substation receives its power from the transmission network, the power is stepped down with a transformer and sent to a bus from which feeders fan out in all directions across the countryside. In an AC system, these feeders carry three-phase power, and tend to follow the major streets near the substation.
   As the distance from the substation grows, the fanout continues as smaller laterals spread out to cover areas missed by the feeders. This tree-like structure grows outward from the substation, but for reliability reasons, usually contains at least one unused backup connection to a nearby substation. This connection can be enabled in case of an emergency, so that a portion of a substation's service territory can be alternatively fed by another substation. This connection can be enabled in case of an emergency, so that a portion of a substation's service territory can be alternatively fed by another substation.
2. **A mesh network topology** - Resembles a web of interconnections, and is thus, more complex than a radial network. In general mesh topologies are applied to transmission of medium voltage distribution grids. A mesh network allows for redundancy. Redundancy allows line failures to occur and power is simply rerouted while lines are deactivated and repaired.

The most common type of transmission grid on the planet at the present is the wide-area synchronous grid (a.k.a., "interconnection" or "synchronous area") is an electrical grid at a regional scale or greater that operates at a synchronized frequency and is electrically tied together during normal system conditions (as a "synchronized zone" at 50Hz or 60Hz). In a synchronous grid all the generators run not only at the same frequency but also at the same phase, each generator maintained by a local governor that regulates the driving torque by controlling the steam supply to the turbine driving it. Generation and consumption must be balanced across the entire grid, because energy is transferred almost instantaneously as it is produced. Energy is stored in the immediate short term by the rotational kinetic energy of the generators. A large failure in one part of the grid - unless quickly compensated for - can cause current to re-route itself to flow from the remaining generators to consumers over transmission lines of insufficient capacity, causing further failures. One downside to a widely connected grid is thus the possibility of cascading failure and widespread power outage. The benefits of synchronous zones include pooling of generation, pooling of load, resulting in significant equalizing effects (i.e., even out the load, reducing generating capacity); common provisioning of reserves. It is not possible to form a wide area synchronous network between two networks operating on different frequency standards (e.g., 50Hz vs. 60hz).

Wide-area synchronous grids can be tied to each other via high-voltage direct current power transmission lines (DC ties), or with variable frequency transformers (VFTs), which permit a controlled flow of energy while also functionally isolating the independent AC frequencies of each side. High-voltage direct current lines or variable frequency transformers can be used to connect two alternating current interconnection networks which are not synchronized with each other. This provides the benefit of interconnection without the need to synchronize an even wider area. For example, compare the wide area synchronous grid map of Europe (above left) with the map of HVDC lines (below right).

6.8.4 The Smart Grid

A 'smart grid' is a type electrical grid that includes variety of operational control and monitoring, and energy "loss", devices. These devices include but are not limited to: "smart" meters, "smart" appliances, renewable energy resources, and energy efficiency resources. Essentially, the "smart grid" is the "grid" enhanced with a variety of control and monitoring devices for improving the efficiency, safety, and reliability of the grid, as well as further increasing the control both industry and the State have over the consumers use of electrical power. With that said, the "smart grid" also gives users more information and tools (when these are made available) to make better choices about their own energy usage.

Here, the term "smart" is used for two purposes. First, it is a marketing term to aid the adoption of these energy control and monitoring devices by consumers and industry -- as in, "it is the smart thing to do pay for these devices, which function [in part] to enhance the monitoring and control of consumers power usage by the power industry". Second, "smart" is similar in meaning to "intelligence", and the term 'intelligence' is applied in engineering to mean that a system is capable of taking decisions or aiding a control system in taking a more informed decision. The addition of these "intelligent" devices to a basic electrical grid adds resiliency to the electric power system by making it better prepared to address emergencies, such as severe storms, earthquakes, large solar flares, and attacks.

6.9 Electrical Power Generation: localization

The localization of electrical power generation can be categorized in three main ways: using network terminology; the presence of a grid connection; and using interconnection-type as a parameter.

There are two network-based categories of generated power localization:

1. **Centralized generation systems** - refers to power which is produced at large generation facilities, and transported though the transmission and distribution grids (far in space) to the end-use.
2. **Distributed generation systems (a.k.a., on-site, decentralized, or localized)** - refers to power that is produced next to (near in space, proximate) its point-of-use. Distributed power generation may also be referred to as on-site generation (OSG), district/decentralized generation, or localized generation. Distributed generators may or may not be connected to a wider transmission and distribution grid. The key criteria in this definition is the proximity to the end-use (and not whether the generators are connected to a wider transmission and distribution grid).

Some generation technologies are more easily distributed than others (e.g., solar panels and wind turbines are relatively easily distributed). Historically, distributed generators were complementary to centralized generation (i.e., they provided solutions to overcome the shortfalls of centralization, such as backup generators for when power was cut to the central generators). Today, however, distributed generators are more widely available because of advances in technology. Conversely, a hydro-electric dam has a definitive position of placement relative to its energy source (the body of water and the dam).

In concern to the presence of a grid connection, a generation system either has a grid connection (is grid connected) or does not have a grid connection (is off grid):

1. **Stand alone (off grid) generation systems** - Systems that are not connected to the grid or do not require the grid. These are the simplest form of electrical power system, with the fewest components. They consist of an electrical source (or several localized and networked sources) and a load(s), which operate independently from the grid. If these system are ever connected to a grid, then their voltage (DC and AC), frequency and waveform (AC) will likely need to be modified to match the grid. Multiple stand alone systems can sometimes be networked.
   - Examples include, backup generators and specific purpose power units. Specific purpose power units are used for applications such as pumping water, electric fences, navigational/safety

signaling, and remote monitoring. These systems are generally designed to run on DC rather than AC (and do not require inverters and control systems).

- Batteries are not required for off grid systems, but their presence has [at least] three benefits:
- Storing energy for use when energy from the primary source is unavailable.
- As a buffer between an intermittent supply and varying/peak demand (a form of load demand management).
- Creating a clean regulated AC supply from an unregulated source.

2. **Grid connected generation systems (a.k.a., grid-joined and grid-tied)** - Systems that are connected to the gird and output power into the grid. In AC grid connected systems the generator voltage and frequency are locked to the grid system, or the voltage and frequency of the generator are modified to match the grid system (this is sometimes known as 'supply regulation'). Also, the generator's output waveform should be a pure sine wave, without harmonics. Some generators are required to reach a minimum speed before they can be connected, so that their output frequency matches the grid frequency. These are also known as: on-grid, grid-tied, utility-interactive, grid-intertied, and grid-direct.
    - When the grid is shut down (for maintenance or emergency), grid connected systems must also be shut down (or disconnected from the grid). Hence, depending upon design, when the grid fails, these systems cannot operate. This is a safety issue. If the grid is shut down and undergoing maintenance, then a grid connected generation system that hasn't been shut down or isn't disconnected could [accidentally] energize the grid and electrocute someone or damage equipment.

NOTE: Islanding *refers to the condition when a portion of the grid becomes temporarily isolated from the main grid but remains energized by its own distributed generation resource(s). Islanding may be unintentional (accidental) or intentional. Unintentional islanding is a potentially hazardous condition, and occurs when a generator fails to properly shut down or disconnect. However, with appropriate safety and control mechanisms, intentional islanding can be used to provide service to mini-grids where the grid is unreliable or parts of the grid have been shut down.*

6.10 Electrical system earthing/grounding

In general, 'earthing' and 'grounding' are different terms for expressing the same concept. The term 'earthing' is more commonly used in some countries and in their accompanying standards, and 'grounding' is more commonly used in other countries and in their own standards. Both terms imply a non-charged state, a common potential, a common point with which the potentials of other points are defined. When this common point is the earth, some standards use the term earth, while other standards use the term ground. When this common point is not the earth, most standards use the word ground, but some still use the word earth.

In electrical circuits, the term 'ground' (or 'earth') can be very confusing, because it has different meanings. The word 'ground' (or 'earth'), without context, could mean any of the following:

1. A 'common' connection, but not connected to Earth.
2. A direct connection to the power supply (usually to the DC negative terminal).
3. A point on a circuit used as a zero-voltage (0V) reference for measuring potential differences (this is the case with most electronics).
4. A connection to the inside of a shielded metal box.
5. A connection to a metal object much larger than the circuit (e.g., car chassis).
6. A connection to a conductive stake driven into the Earth (or a connection to a metal water pipe which extends out of a building into the earth). In an electrical power system, the ground or earth is a conductor that provides a low impedance path to the earth to prevent hazardous voltages from appearing on equipment.

In general, grounding (uncharging) is the process of removing the excess charge on an object by means of the transfer of electric charge (electrons) between it and another object of substantial size. When a charged object is "grounded", the excess charge is balanced by the transfer of electrons between the charged object and a ground. A ground is simply an object that serves as a seemingly infinite reservoir of electrons; the ground is capable of transferring electrons to or receiving electrons from a charged object in order to neutralize the charge on that object. Grounding requires a conducting pathway. "Ground" may be used as a reference point for measurement. The "earth" is the most common ground reference. It is sometimes said that 'earth'/'ground' is a statement of voltage.

WATER ANALOGY: *Imagine a lake, either man-made or natural, then the top of the lake is equivalent to ground, a place where the water/charge is all at the same potential and where lots of flow/current can easily go in or out without changing the potential.*

In an electrical circuit, ground or earth is the reference point from which voltages are measured, a common return path for electric current, or a direct physical connection to the Earth. In electrical power distribution systems, a protective ground conductor is an essential part of the safety earthing/ground system. Here, 'earth'/'ground' refers to a body that has such a large charge sink/source capacity that for circuit purposes any current flows do not affect its potential. In electrical systems, the Earth is commonly used as ground because it is very large and conductive (generally); it also then serves as a common reference point. The minerals and moisture in the Earth (in soil) will conduct.

HISTORICAL NOTE: *An essential part of radio is an antenna, and an essential part of early antennas was a connection to the Earth.*

The concept of system grounding is extremely important, as it affects the susceptibility of the system to voltage transients, determines the types of loads the system can accommodate, and helps to determine the system protection requirements.

NOTE: *In a DC circuit, current from a battery leaves the positive terminal and it has to return to the negative terminal before any current can flow. So connecting it to ground has no effect although it also won't do any harm. The circuit will operate just the same if you connect one side of the battery, while the circuit is complete, to ground.*

An **earthing system** (**grounding system**) is circuitry which connects parts of an electric circuit with the ground (earth), but not necessarily the Earth, thus defining the electric potential circuit. If a fault within an electrical device connects a live supply conductor to an exposed conductive surface, anyone touching it while electrically connected to the earth will complete a circuit back to the earthed supply conductor and receive an electric shock.

A **protective earth (PE) connection (a.k.a., equipment grounding conductor)** avoids electrical shocks by keeping the exposed conductive surfaces of a device at earth potential. To avoid possible voltage drop no current is allowed to flow in this conductor under normal circumstances. In the event of a fault, currents will flow that should trip or blow the fuse or circuit breaker protecting the circuit. A high impedance line-to-ground fault insufficient to trip the overcurrent protection may still trip a residual-current device (ground fault circuit interrupter or GFCI) if one is present. This disconnection in the event of a dangerous condition before someone receives a shock, is a fundamental tenet of best practice wiring, and is often referred to as automatic disconnection of supply (ADS). The alternative is 'defence in depth', where multiple independent failures must occur to expose a dangerous condition - reinforced or double insulation come into this latter category.

In contrast to protective earth (PE), a **functional earth connection (functional ground connection)** serves a purpose other than shock protection, and may carry power or signal current as part of normal operation. The most important example of a functional earth is the neutral line in an AC electrical power supply system. It is a current-carrying conductor connected to earth, often, but not always, at only one point to avoid flow of currents through the earth. This connection is sometimes called a "grounded supply conductor" to distinguish it from the "equipment grounding conductor". Common examples of devices that use functional earth/ground connections include surge suppressors, electromagnetic interference filters, certain antennas, and measurement instruments. Great care must be taken when functional earth's from different systems meet to avoid unwanted and possibly dangerous interactions, for example lightning conductors and telecom systems must only be connected in a way that cannot cause the energy of the lightning strike to be redirected into the telecom network.

Earthing/grounding systems can be subdivided at a top-level into low-voltage earthing/grounding systems and high voltage earthing/grounding systems.:

1. In low-voltage distribution networks, which distribute the electric power to the widest class of end users, the main concern for design of earthing systems is safety of consumers who use the electric appliances and their protection against electric shocks. The earthing system, in combination with protective devices such as fuses and residual current devices, must ultimately ensure that a person must not come into touch with a metallic object whose potential relative to the person's potential exceeds a "safe" threshold, typically set at about 50 V.
2. In high-voltage networks (above 1 kV), which are far less accessible to the "general population", the focus of earthing system design is less on safety and more on reliability of supply, reliability of protection, and impact on the equipment in presence of a short circuit. Only the magnitude of phase-to-ground short circuits, which are the most common, is significantly affected with the choice of earthing system, as the current path is mostly closed through the earth. Three-phase HV/MV power transformers, located in distribution substations, are the most common source of supply for distribution networks, and type of grounding of their neutral determines the earthing system.

NOTE: *A connection to the earth/ground is essential to protect a structure from lightning strikes. It directs the lightning through the earthing system and into the ground rod rather than passing through the structure.*

In an electrical current distribution system, there are three possible elements through which current may travel:

1. **The positive wire (DC)** - current enters or current returns (depends on terminology).
2. **The negative wire (DC)** - current enters or current returns (depends on terminology).
3. **The hot wire (AC, positive wire)** is the path for current to flow from source to load.
4. **The neutral wire (AC)** is the return path for the current from the load; it is the return path provided to complete circuit. Neutral carries current equal to that carried by the Hot wire. It is the return path for the Hot. Things in nature like to be "balanced". Without a return path, there's no movement of electrons, and thus, no "electricity". In a single phase branch circuit, the current on the hot wire and neutral should be identical (unless there is ground leakage).
5. **The ground/earthed wire (AC & DC)** is a low impedance pathway between things that might become energized (i.e., has voltage), but are not supposed to be. Under normal conditions, a grounding conductor does not carry current (or voltage). If a fault occurs and if ground is energized/connected, then it completes the circuit back to the source. In other words, ground is there for safety, and it should not carry current, except when something has failed.

NOTE: *When a connection has not been made between the neutral point and earth/ground, it is said that the neutral is unearthed/ungrounded.*

6.10.1 AC voltage specific issues

An AC voltage system will pass current in the following ways:

1. Balanced: If a three phase load is balanced, and also if the generator system is perfectly balanced then equal current flows through all three wires and no current flows through the neutral line. When the electrical system is "balanced" (a balanced load), the neutral line/wire is 0V (no current flow).
2. Unbalanced: When the system is "unbalanced", then the current through neutral will not be 0. In all the cases (except faults), no current flows through ground wire.
3. Nominal single phase operation (2 wire): In the case of single phase AC (with 2 wires; not SWER), the return path is the neutral wire. Here the question of balanced or unbalanced doesn't arise.
4. Nominal single-wire earth return (SWER; 1 wire): In the case of single-phase earth return, one conducting wire passes from the source to the load, and the grounded/earthed wire is the neutral return path from the load.
5. Except in the case of SWER, only in case of a fault current will flow through the earth wire, otherwise no current flows to "ground".

In electrical AC transmission/distribution systems, neutral sometimes goes "to ground" (to earth), and sometimes does not go "to ground" (to earth). The United States National Electric Code (US NFPA 70) requires that neutral and ground be bonded at the main service entrance for residential electrical service. This bonding is done in this and only this location. Bonding ground and neutral again elsewhere in the system will create parallel ground paths, which is very dangerous.

In AC (more than single phase), ground is connected to neutral for safety. If the bond between ground and neutral is removed, the system will have a "floating neutral", that is, a neutral that has no reference to earth ground. However, in such a configuration, when a ground fault occurs (a specific type of short where the hot wire touches something grounded), it will not trip the breaker. This is a safety hazard, because when a ground fault does occur, no one will know (because breaker isn't tripped). Therein, everything that is grounded may be energized up to system voltage. Hence, ground and neutral are connected at the main service entrance so that when a ground fault occurs, the breaker is tripped, and power is cut to the circuit. It's a safety issue, with a minor secondary issue being improperly grounded equipment can be prone to premature failure.

Some AC grids, like the United States and New Zealand grids, allow electric charges (electricity) to be released back (shunted) into the Earth. AC grids in other countries do not allow this and add an additional wire to the transmission/distribution system as the return path. One name for an electrical transmission/distribution system that continuously releases charge into Earth is called single-wire earth return.

Single-wire earth return (SWER) or **single-wire ground** return is a single-wire transmission line which supplies single-phase AC voltage electric power from an electrical power grid -- an electrical transmission/distribution method using only one conductor with the return path through earth. Single-wire earth return systems are significantly different from the three-phase, three-wire and single-phase, two-wire systems. As the name implies, it is a single-wire distribution system in which all equipment is grounded to earth and the load current returns through the earth. Its loads are light and its lines are long, often causing the current to have a leading power factor. Its distinguishing feature is that the earth (or sometimes a body of water) is used as the return path for the current, to avoid the need for a second wire (or neutral wire) to act as a return path. The SWER line is a single conductor that may stretch for tens or even hundreds of kilometres, with a number of distribution transformers along its length. At each transformer, such as a customer's premises, current flows from the line, through the primary coil of a step-down isolation transformer, to earth through an earth stake. From the earth stake, it is claimed by some that "the current eventually finds its way back to the main step-up transformer at the head of the line, completing the circuit". SWER is therefore a practical example of a 'phantom loop'.

There are several issues with single-wire earth return, including but not limited to:

1. The SWER must be designed to prevent dangerous step and touch potentials.
2. Telephone interference, similar to 2 wire single phase lines, worse than three-phase lines.
3. Load balance problems can erode efficiency.
4. Load density limitations.
5. Voltage control can be difficult.
6. Power quality can be compromised.
7. The Earth in a given location is an inadequate composition to conduct electricity.
8. Stray voltage and interference with the Earth's natural electrical currents.

6.10.2 Topological layouts for grounding/ earth systems

There are several basic topological layouts for a grounding/earthing system: ungrounded; solidly grounded; impedance/resistance grounded; and reactance grounded.

6.10.2.1 Ungrounded

Electrical power systems that are operated with no intentional connection to earth ground are described as ungrounded. The term "ungrounded system" is actually a misnomer, since every system is grounded through its inherent charging capacitance to ground.

Some systems should not be grounded. A system may be left ungrounded when it is determined that the hazards of grounding outweigh the safety benefits of grounding. One such determined system-type may be an isolated hospital operating room power system, which is a local distribution power system of limited size. Such a system will be left ungrounded, because it is considered unacceptable to have a power outage during a surgical procedure.

Advantages include:

1. Offers a low value of current flow for line-to-line ground fault (5A or less).
2. Presents no flash hazard to personnel for accidental line-to-ground fault.
3. Assures continued operation of processes on the first occurrence of a line-to-ground fault.
4. Low probability of line-to-ground arcing fault escalating to phase-to-phase or 3-phase fault.

Disadvantages include:

1. Difficult to locate line-to-ground fault.
2. Doesn't control transient over voltages.
3. Cost of system maintenance is higher due to labor involved in locating ground faults.
4. A second ground fault on another phase will result in a phase-to-phase short circuit.

6.10.2.2 Solidly grounded

Grounding conductors are connected to earth ground with no intentional added impedance in the circuit. A main secondary circuit breaker is a vital component required in this system, although it has no bearing in other grounding systems. This component is large in size because it has to carry the full load current of the transformer. Back-up generators are frequently used in this type of grounding system in case a fault shuts down a production process. A solidly grounded system has high values of current ranging between 10kA and 20kA.

Advantages include:

1. Good control of transient over-voltage from neutral to ground.
2. Allows user to easily locate faults.
3. Can supply line-neutral loads.

Disadvantages include:

1. Poses severe arc flash hazards.
2. Requires the purchase and installation of an expensive main breaker.
3. Unplanned interruption of production process.
4. Potential for severe equipment damage during a fault.
5. High values of fault current.
6. Likely escalation of single-phase fault to 3-phase fault.
7. Creates problems on the primary system.

6.10.2.3 Impedance grounded system

An impedance/resistance grounded system incorporates the benefits of both the grounded and the ungrounded system. Impedance grounded systems include high resistance ground (HRG) and low resistance ground (LRG) configurations. Low-resistance grounding (impedance-type) is typically used in medium voltage systems, which have only 3-wire loads, such as motors, where limiting damage to the equipment during a ground fault is important enough to include the resistor, but it is acceptable to take the system offline for a ground fault.

1. **High-resistance grounding (impedance-type)** systems are commonly used in plants and mills where continued operation of processes is paramount in the event of a fault. High-resistance

grounding is normally accomplished by connecting the high side of a single-phase distribution transformer between the system neutral and ground, and connecting a resistor across the low-voltage secondary to provide the desired lower value of high side ground current. With an HRG system, service is maintained even during a ground fault condition. If a fault does occur, control systems can locate and correct the fault, or shut down the system in a safe and orderly manner. An HRG system limits ground fault current to between 1A and 10A.

Advantages include:

1. Limits the ground fault current to a low level.
2. Reduces electric shock hazards.
3. Controls transient over voltages.
4. Reduces the mechanical stresses in circuits and equipment.
5. Maintains continuity of service.
6. Reduces the line voltage drop caused by the occurrence and clearing of a ground fault.

Disadvantages include:

1. High frequencies can appear as nuisance alarms.
2. Ground fault may be left on system for an extended period of time.

6.10.2.4 Reactance grounded

Reactance grounded (reactance-grounded) describes the case in which a reactor is connected between the system neutral and ground. it is commonly used in the neutrals of large AC generators. It provides limiting effect to fault current passage through the circuit and also doesn't consume active power.

6.10.3 Fault types

In electrical power systems, a fault or fault current is any abnormal electric current. The design of systems to detect and interrupt power system faults is the main objective of power-system protection. There are four principal categories of fault:

1. A short circuit is a fault in which current bypasses the normal load.
2. An open-circuit fault occurs if a circuit is interrupted by some failure.
3. Phase faults: In three-phase systems, a fault may involve one or more phases and ground, or may occur only between phases. In a "ground fault" or "earth fault", current flows into the earth. In a polyphase system, a fault may affect all phases equally which is a "symmetrical fault".
4. In a "ground fault" or "earth fault", current flows into the earth.

There are several types of faults that an electrical system must be designed to withstand. Electrical equipment is typically sized and noted with a fault current rating based on fault calculations. A designer must account for the worst-case scenario. Among the four principal types of faults are several categories of fault:

1. **Transient fault** - a fault that is no longer present if power is disconnected for a short time and then restored; or an insulation fault which only temporarily affects a device's dielectric properties which are restored after a short time
2. **Persistent fault** - a fault that does not disappear when power is disconnected.
3. **Symmetric fault (balanced fault)** - a fault in an AC phased system that affects each of the [three] phases equally.
4. **Asymmetric fault (unbalanced fault)** - a fault in an AC phased system that does not affect each of the [three] phases equally. Common asymmetric faults include:
   A. Line-to-line (line-line) - a short circuit between lines, caused by ionization of air, or when lines come into physical contact, for example due to a broken insulator.
   B. Line-to-ground - a short circuit between one line and ground.
   C. Double line-to-ground - two lines come into contact with the ground (and each other).
5. **Bolted fault** - a fault with zero impedance, giving the maximum prospective short-circuit current.
6. **Arcing fault** - a fault where the system voltage is high enough that an electric arc may form between power system conductors and ground. Arcing faults are often formed via intermittent failures between phases or phase-to-ground. They're discontinuous currents that alternately strike, extinguish, and strike again.

7 Combustion power systems

CLARIFICATION: *In practice, combustion and burning refer to the chemical reaction's occurrence in different environments: combustion refers to when the process occurs in an environment with a "fixed" amount of air/oxygen, whereas burning refers to when the process has access to an infinite amount of air/oxygen (i.e., done in the "open"). And, fire is said to exist when the surrounding atmospheric air is the source of the oxidant/oxygen.*

Thermal radiation, process by which energy, in the form of electromagnetic radiation, is emitted by a heated surface in all directions and travels directly to its point of absorption at the speed of light; thermal radiation does not require an intervening medium to carry it.

Chemical kinetics is the quantitative study of chemical systems that are changing with time. (Thermodynamics, another of the major branches of physical chemistry, applies to systems at equilibrium—those that do not change with time.) Chemical kinetics, the branch of physical chemistry that is concerned with understanding the rates of chemical reactions. It is to be contrasted with thermodynamics, which deals with the direction in which a process occurs but in itself tells nothing about its rate. Thermodynamics is time's arrow, while chemical kinetics is time's clock.

Combustion, **burning,** and **firing** refer to the same complex sequence of exothermic redox chemical reactions between a fuel and an oxidant accompanied by the production of heat, or both heat and light. In other words, combustion is an oxidative decomposition in which oxygen (the oxidant) oxidizes a fuel. Combustion is the rapid oxidation of a material (the fuel), which releases energy (as, at least, heat). As a chemical reaction, combustion is an irreversible process, and leads to the formation of principal gaseous products (i.e., new chemical species). The interdisciplinary scientific study of combustion combines [at least] heat transfer, thermodynamics, chemical kinetics, and multiphase turbulent fluid flow. Because the reaction is exothermic, the product gases heat up and expand, and in turn can be harnessed to do work. Hence, combustion produces the following sources/carriers of energy:

1. Heat (heat energy transfer; thermal EM radiation/power)
2. Light (electromagnetic radiation; EM power)
3. Turbulent fluid flow (directly, fluid power; indirectly, mechanical power)

All forms of combustion involve the redox chemical reaction of oxidation and reduction, which control the release of heat (light and fluid flow) from the chemical reaction between a fuel and an oxidizer. Therein, combustion is a sequence of elementary radical reactions. For instance, most solid fuels first undergo endothermic pyrolysis (thermochemical decomposition) to produce gaseous fuels whose combustion then supplies the heat required to produce more of them.

Combustion reactions are exothermic (i.e., they give off thermal energy, heat emitting). In a combustion reaction, the fuel (substance undergoing combustion) is oxidized by the oxidant. Usually the oxidizing agent is molecular oxygen (O_2), but there are other oxidants (e.g., halogens; hydrogen burns in chlorine). The amount of heat released in a chemical reaction can be calculated through thermodynamics.

NOTE: *During chemical reactions, energy is either released to the environment (exothermic reaction) or absorbed from the environment (endothermic reaction). In other words, a chemical reaction that releases energy is termed exothermic, and one that absorbs energy is termed endothermic.*

The products (substances produced by the reaction) of combustion will always have a higher oxidation state than the reactants (substances that start the reaction). In combustion reactions, heat, light, and fluid flow are produced, and work can be done from the transfer of energy. However, for oxidation reactions without combustion, this is not always true.

NOTE: *Carbon is the universal element of organic compounds. The molecule of an organic substance must have at least one carbon atom in its molecule. Notice that water does not contain any carbon atom in its molecule, H_2O. Hence, water is only an inorganic compound.*

Both organic and inorganic (e.g., gunpowder and magnesium) compounds are capable of combustion, whereupon they become oxidized. All organic matter can be combusted, but only some inorganic matter can be combusted. The burning of a combustible substance can occur in gaseous, liquid, or solid form.

NOTE: *Substances that are able to combust [under useful conditions] are called* **flammable**.

When organic molecules combust, the reaction products are [at least] carbon dioxide (CO_2) and water (H_2O); however, the products will vary depending upon the starting material. In the process of burning, the carbon in these organic fuel substances becomes bonded with oxygen, while some of the oxygen used to "burn" the fuel bonds to the hydrogen atoms from the fuel. Combustion reactions are good examples of **redox reactions** where one molecule gains oxygen (is oxidized) and one molecule gains hydrogen (is reduced).

NOTE: *Fires occur naturally, ignited by lightning strikes, significant static electricity, or by volcanic products.*

In order to visualize combustion as a chemical process, imagine separating the carbon and hydrogen atoms of a hydrocarbon molecule (e.g., alkane) and the oxygen atoms of oxygen molecules, and letting the individual atoms attract to form carbon dioxide and water. Separating the atoms (i.e., breaking bonds) involves the input of energy (bond dissociation energy), because the bonding electrons represent a negative charge density, which attracts and holds the positive nuclei of bonded atoms together. As the atoms move in separation, kinetic energy arises (as exothermic energy) from the potential energy of the bonds. When their attracted connection is complete, a new potential energy state/well is formed (e.g., carbon-oxygen as CO_2 and hydrogen-oxygen as H_2O bonds).

> **NOTE:** *Electronegativity, symbol χ, is a chemical property that describes the tendency of an atom or a functional group to attract electrons (or electron density) towards itself. An atom's electronegativity is affected by both its atomic number and the distance at which its valence electrons reside from the charged nucleus. The higher the associated electronegativity number, the more an element or compound attracts electrons towards it. The opposite of electronegativity is electropositivity: a measure of an element's ability to donate electrons.*

As a general rule, the greater the electronegativity difference between bonded atoms, the stronger the bonds. Hence, during combustion, relatively weak, low-electronegativity-difference bonds (carbon-carbon, carbon-hydrogen, and oxygen-oxygen) are replaced by stronger, high-electronegativity-difference bonds (carbon-oxygen and hydrogen-oxygen). Oxygen is highly electronegative, and in the presence of sufficient [endothermic] energy input will pull the bonded electrons of other atoms toward itself, separating them and reconnecting in favour of a gain in an electron.

7.1 Oxidation reduction reactions (Redox)

In general, combustion is considered a redox reaction. Oxidation reduction reactions (a.k.a., redox reactions) are a basic and common type of chemical reaction found in nature. In a redox reaction, two reactions occur: oxidation and reduction. Take note that an oxidation reaction cannot happen without a corresponding reduction reaction.

Originally oxidation reactions were identified as reactions in which oxygen gas participates, which is why this type of reaction is presently known as oxidation (e.g., $2H_2 + O_2 \rightarrow 2H_2O$). However, as scientific understanding grew, it was discovered that another way to characterize oxidation is through a loss of hydrogen, and not the adding of oxygen (e.g., $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$). As scientific understanding grew further, it was discovered that another way to characterize oxidation is through the loss of electrons, where there is no observation of an oxide formation or hydrogen loss (e.g., $Mg + Cl_2 \rightarrow Mg^{2+} + 2Cl^-$). Therefore, today, oxidation means either:

1. Gain of oxygen
2. Loss of hydrogen
3. Loss of electrons

Characterizations 2 and 3 are not combustion reactions. And, and not all observations of characterization 1 are combustion reactions. Therefore, combustion reactions are oxidation reactions, but not all oxidation reactions are combustion reactions. The term oxidation is now something of a misnomer, because it starts with the same prefix as oxygen, suggesting that oxygen is involved in the process, which is not always true and only representative of one characterization of the process.

Today, scientific understanding has determined that a more appropriate definition of oxidation and reduction includes the movement of electrons between the compounds involved. Hence, oxidation-reduction reactions or redox reactions are reactions in which electrons are transferred from one atom or molecule to another. In the most broad definition of redox reactions:

1. Oxidation is the loss of electrons (or an increase in oxidation state by a molecule, atom, or ion).
2. Reduction is the gain of electrons (or a decrease in oxidation state by a molecule, atom, or ion).

In terms of combustion, whenever combustion takes place, oxidation is the end result via redox reaction. For combustion, the usual oxidant is oxygen, but for an oxidation reaction to take place, oxygen is not essential -- all combustion is oxidation, but oxidation includes other reactions.

> **NOTE:** *Rust[ing] is another example of a redox reaction. When something made of iron is exposed to oxygen atoms in the presence of moisture, the iron acquires electrons from the oxygen. The iron and oxygen have opposite charges, which attract, converting the iron into a flaky reddish material called iron oxide (rust).*

7.2 Elements of the combustion process

All forms of combustion require at least the following three primary inputs and/or conditions, and if any of these are removed the combustion (fire) will cease to exist:

1. **Fuel (reactant):** The input that burns - a combustible/flammable material. The fuel can be a solid, liquid, or gas that is capable of undergoing combustion. The simplest possible fuel is pure hydrogen gas. During combustion, the fuel donates electrons. The lower a substances electronegativity, the more reactive it will be as a fuel.
2. **Oxidizer (reactant):** The molecule that accepts electrons, which is then reduced. Combustion

requires the fuel to be oxidized, that is, it donates electrons. The oxidizer must be of sufficient quantity to support combustion. Oxygen is a good oxidizer, because it is so electronegative, which means it will relatively easily accept electrons. Only fluorine is more electronegative than oxygen; however, it rarely exists in free elementary form. The predominant oxidizer used in most manufacturing heating processes is atmospheric air.

3. **Heat (i.e., energy input):** Sufficient heat to bring the fuel to its ignition temperature and keep it there. In a combustion reaction, there is energy input and energy output. The input energy (power) starts and/or ignites the reaction, which is true of most chemical reactions. In other words, in order for combustion to occur, there must be sufficient [thermal] energy to bring the fuel (in the presence of an oxidizer) to its ignition temperature, and keep it there. Note that this so-called "activation energy" is usually much less than the energy ultimately released from combustion (as energy output). Metaphorically, ignition/input energy is like rolling a boulder some distance in order to let the natural process of "falling" take over as it rolls it down a hill; as it begins falling down the hill it releases ["gravitational"] potential energy.
4. **Chemical chain reaction (redox)** - In other words, the redox reaction. The heat produced by combustion can make the reaction self-sustaining. Combustion (burning and fires) start when a flammable or a combustible material, in combination with a sufficient quantity of an oxidizer, such as oxygen gas or another oxygen-rich compound (though non-oxygen oxidizers exist), is exposed to a source of heat (or ambient temperature above the flash point for the fuel/oxidizer mix), and is able to sustain a rate of rapid oxidation that produces a chemical chain reaction.

NOTE: *Fire (combustion) is normally represented as a triangle of only three inputs: oxygen; heat; and fuel. However, it is more accurate to model fire as a combination of four elements, because fire can be extinguished by removing any one of these four (and not just three) elementary conditional inputs.*

7.3 The Combustion continuum (types of combustion)

The terms used to describe combustive decomposition depends on characteristics, such as the speed of the reaction. Combustion can be divided into several types:

7.3.3.1 Unintentional combustion

- **Spontaneous combustion** - Combustion in which substances suddenly burst into flames, without the application of any apparent or intentional cause.

7.3.3.2 A type of combustion, but also, part of the definition of combustion

- **Rapid combustion** - Combustion in which substances burn rapidly to produce heat and light. Combustion, itself, is generally defined as a rapid redox reaction.

7.3.3.3 The continuum from a slower speed of reaction to a faster speed of reaction

1. **Smouldering** - The slow, low-temperature, flameless form of combustion, sustained by the heat evolved when oxygen directly attacks the surface of a condensed-phase fuel. Smouldering is typically an incomplete combustion reaction.
2. **Deflagration (a.k.a., mild burn)** - The opposite of an explosion is termed a mild burn, where the intended products of combustion (such as CO2, H2O, and N2) eventually dominate the composition, as exists in most well-controlled combustion processes (car engine, jet engine, furnace, gas stove top, fireplace, power plant boiler, etc). Deflagration results in subsonic flame velocities. A combustive reaction occurs at less than the speed of sound, it is called deflagration. This term can also be applied to what we mean by the term "burning," in which the flame speed is less than the speed of sound.
3. **Detonation/explosion** - A detonation combustion results in a shock wave of supersonic velocities and can loosely be described as an explosion. Detonation is similar to explosion with the difference lying in the fact that the rapid increase in volume is so high that the production of a supersonic shock wave takes place. The detonation of an explosive (fuel) causes a reaction front that moves faster than the speed of sound (~ 741 mph, or 331 m/s). Take note that there are non-combustive forms of explosion. Technically, it is scientifically incorrect to define an explosion as "a type of combustion". Explosions are defined primarily as a rapid, violent uncontrolled release of energy. Although explosions often involve some kind of temperature difference, they quite commonly occur though means that are totally without combustion - such as mechanical explosions driven by gases. It is sometimes said that combustion means burning, and explosion

means bursting.

7.3.3.4 A combustion-related process, but technically not a type of combustion

- **Flame/glowing** - Some combustion processes and combusting substances produce a flame. Substances which vaporise while burning give flame and those which do not vaporise while burning do not give flame. Fuels that burn with a flame produce light. Flames and/or glowing represent combustion reactions that are propagating through space at subsonic velocity, and are accompanied by the emission of heat and light (EM radiation). In other words, any chemical process that produces light and heat as either glow or flames is combustion. What we observe as fire is only a small portion of the combustion/burning reaction called the 'flame'. The flame is the result of complex interactions of chemical and physical processes. The flame is the part of the fire made of burning gaseous compounds and fine suspended particles. Evaporation/gasification of the fuels (if these are liquid/solid) and subsequent thermal degradation into smaller molecules and/or reactive radical species forms the gaseous compounds. The composition of the flame can change depending on the nature of the fuel. In their material composition, flames are mostly made up of reaction by-products, such as carbon dioxide (carbon and oxygen), water (hydrogen and oxygen, and oxygen. A flame is a mixture of reacting gases and solids emitting visible, infrared, and sometimes ultraviolet light, the frequency spectrum of which depends on the chemical composition of the burning material and intermediate reaction products. Because flames emit energy in the form of light, the flame is referred to as the visible part of the fire. The color of a flame depends on a variety of conditions; temperature, chemical composition, and the amount of oxygen present can change the color of a flame/fire. Any smoke emitted from the yellow flame is unburnt fuel, also called soot. The more smoke and the more yellow the flame, the more "impure" it is, such that the reactants haven't fully combusted.

7.4 The Two reactant forms of combustion

NOTE: *Combustion is used either directly or indirectly to produce virtually every product in common use. Combustion processes produce and refine fuel, generate electricity and other forms of power, prepare foods and other materials, and transport goods.*

The degree of combustion can be measured and analyzed with test equipment. Combustion analyzers may be used to test the efficiency of a burner during the combustion process. In general, there are considered to be two forms of combustion: complete combustion and incomplete combustion:

1. **Complete combustion** - occurs when a sufficient supply of the oxidant is present so that the elements in the fuel react fully - complete burning of the fuel. In complete combustion, the reactant burns to the extent that it produces no (or, a limited number of) by-products. When a fuel undergoes complete combustion, it releases the maximum amount of energy from the fuel being reacted. Complete combustion is usually characterized by a blue flame. In other words, a more complete combustion of gas, for example, has a dim blue color due to the emission of single-wavelength radiation from various electron transitions in the excited molecules formed in the flame. For example, the complete combustion of a ideally pure hydrocarbon with oxygen would produce only carbon dioxide and water (i.e., hydrocarbon + oxygen > carbon dioxide + water).
2. **Incomplete combustion** - occurs when a insufficient supply of the oxidant is present so that the elements in the fuel are not fully reacted - incomplete burning of the fuel. In incomplete combustion, the reactant burns to the extent that it produces by-products. Incomplete combustion is often undesirable because it releases less energy than complete combustion and produces carbon monoxide which is a poisonous gas. Incomplete combustion will produce pure carbon (soot), which is "messy" and will build up on/in equipment. Incomplete combustion is characterized by an orange coloured flame. In incomplete combustion, products of pyrolysis remain unburnt and contaminate the smoke with noxious particulate matter and gases. Partially oxidized compounds may also be present, and are often toxic. When hydrocarbon fuels are used, the products after a complete burning are usually carbon dioxide and water. However, if the burning didn't happen completely, carbon monoxide and other particles can be released into the atmosphere as pollution. In other words, the incomplete combustion of a ideally pure hydrocarbon with oxygen would produce carbon monoxide, carbon and water, as well as carbon dioxide (i.e., hydrocarbon + oxygen > carbon monoxide + carbon + water, and carbon dioxide). The quality of combustion can be improved by selecting a purer fuel and/or improving the designs of combustion devices.

Note that any combustion at high temperatures in atmospheric air, which is 78 percent nitrogen, will also create small amounts of several nitrogen oxides, commonly referred to as NO_x, since the combustion of nitrogen is thermodynamically favored at high, but not low temperatures. Since combustion is rarely clean, flue gas cleaning or catalytic converters may be required. Further improvement of combustion outputs is achievable by catalytic after-burning devices (such as catalytic converters) or by the simple partial return of the exhaust gases into the combustion process.

NOTE: *When physical elements are burned, the products are primarily the most common oxides. Carbon will yield carbon dioxide, sulfur will yield sulfur dioxide, and iron will yield iron(III) oxide.*

Combustion that produces more relative by-products is dangerous to the health of biological organisms, uses more fuel, and leaves more residue when the chemical reaction is finished. The more efficient and cleaner a combustion, the easier it is to work with and the better it is for practical applications. A more optimal fire is one that uses less fuel and leaves less by-products.

7.5 Fuel

Substances or materials which undergo combustion/burning/firing are known as fuels. A fuel is any substance (material) capable of undergoing combustion and transferring energy in the form of heat, or heat and light. A fuel is that which is flammable. Combustion, burning, and/or setting fuel aflame will release usable energy. Every phase of matter may be formed into, or otherwise compose, a fuel: solid fuel; liquid fuel; gaseous fuel; plasma fuel. Fuels can be used either by themselves, or they can be mixed with other fuels (into a 'fuel mixture'). Different fuels produce different amounts of heat and light, and different by-products when combusted.

NOTE: *Technically, only vapors burn, not liquids or solids. Each type of fuel has a different volatility. Volatility is a measure of how rapidly the liquid turns into vapors. The vapors still must be raised to at least its flash point before ignition can occur.*

Fuels can generally be classified as gaseous, liquid, or solid. In cases where a solid fuel is finely ground, such as pulverized coal, and can be transported in an air stream, its control characteristics approach those of a gaseous fuel. Liquid fuels, as they are atomized and sprayed into a furnace, also have control characteristics similar to those of a gaseous fuel. The control treatment of a solid fuel that is not finely ground is quite different from that of a gaseous or liquid fuel.

NOTE: *A chemical analysis of the fuel will assist in determining how much air (oxidizer) must be mixed with it for complete combustion. The relationship between fuel and air is called the 'fuel/air ratio' or 'fuel/air mixture'. In an engine-type combustion system, the mixture is typically adjusted by controlling the amount of fuel or air entering a carburettor. Therein, supplying too much fuel is called a "rich" mixture and causes excess emissions or smoke from the exhaust. Supplying too little fuel is called a "lean" mixture and causes poor heat generation and a rough running engine.*

The choice of fuel has an important influence on a combustion system and its heat transfer ability. In general, solid fuels (e.g., coal and liquid fuels, like oil) produce luminous flames when combusted, which contain soot particles that radiate like blackbodies to the heat load. Gaseous fuels (e.g., natural gas) often produce non-luminous flames, because they burn more cleanly and completely, and are less likely to produce soot particles. A fuel like hydrogen is completely non-luminous, because there is no carbon available to produce soot. (Londerville, 2013)

There are combustion situations where highly radiant flames are required, and therein, a luminous flame is preferred. Alternatively, in cases where only convection heat transfer is applicable, then a non-luminous flame may be preferable in order to minimize the possibility of contaminating the heat load with soot particles from a luminous flame. (Londerville, 2013)

All flammable material has a flash point and an ignition point:

1. The **flash point** of fuel is the lowest temperature at which sufficient vapors are given off for in a momentary flash when an ignition source is applied near the surface. Flash point is the minimum temperature at which liquid will give off vapours that will ignite. The fuel does not have to remain ignited, and may just "flash".
2. The **ignition point/temperature (a.k.a., auto-ignition)** is the temperature at which the ignited material provides enough heat to start combustion (i.e., start burning). Ignition temperature is the minimum temperature that a substance must be raised to before it will ignite. The "ignition temperature" is the temperature that will start a fuel to rapidly ignite with an oxidizer causing combustion to take place. In other words, it is the lowest temperature at which a combustible substance, when heated catches fire and continues to burn.
3. **Fire point** is the lowest temperature at which a fuel will give off vapours sufficient to cause self sustained combustion for 5 seconds or more. Fire point is nothing but the minimum or the lowest temperature at which vaporization occurs, and

these vapors will start to burn (and burn for at least 5 seconds or more), provided an external source of ignition.

CLARIFICATION - Ignition *is the process/phenomena of initiating the overall burning/combustion/firing process. Before a substance will burn, it must be heated to its ignition point, or kindling temperature. In other words, sufficient heat must be present to ignite the combustion process. The 'ignition temperature' is the temperature that will start a fuel to rapidly ignite in the presence of an oxidizer (e.g., oxygen) causing combustion to take place. Regardless of the fuel, it must be vaporized in order to burn. Oil, a liquid, and coal, a solid, must be heated to the point where gaseous vapors are rapidly given off. It's these vapors which burn, NOT the solid or liquid. This is what makes it possible, for example, to put out a match in a bucket of light oil that is below its flash point.*

7.5.1 Material sources of fuel for combustion

Besides the different phases of matter, there are several different sources of fuel for useful power-oriented combustion purposes. Take not that all of these sources of fuel originate from (or are themselves) biomass, and hence, from solar radiation and photosynthetic production.

1. **Biomass** - Raw biomass from recently living organisms.
2. **Biofuel -** A biofuel is a fuel that is produced through contemporary biological processes, such as agriculture and anaerobic digestion, rather than a fuel produced by geological processes such as those involved in the formation of fossil fuels, such as coal and petroleum, from prehistoric biological matter. A biofuel is a fuel, produced from living organisms.
3. **Fossil fuel** - A fossil fuel is a highly concentrated store of ancient photosynthetic production. Fossil fuel is a general term for buried combustible geologic deposits of organic materials, formed from decayed plants and animals (and their excrement) that have been converted to its three primary forms (as crude and heavy oils, coal, or natural gas) by exposure to heat and pressure in the earth's crust (geological activity) over thousands to millions of years. As organic material (former living organisms) decay within the earth they decompose into hydrocarbon-type material. And, because of their molecular makeup (i.e., primarily carbon and hydrogen), they readily combine with oxygen under combustion to produce a different compound and release heat. After organisms die, their organic material settles to the surface of the planet, which over time, becomes more deeply buried. This buried organic material eventually forms a layer of partly decomposed spongy-like material called 'peat'. Peat, itself, may be used as a soil nutrient amendment or a fuel. As the peat subsides further and is exposed to greater heat and pressure, it forms into the various types of fossil fuels (i.e., underground hydrocarbon resources). In other words, 'peat' is the material precursor to all fossil fuels.
4. **Refined fossil fuels** - In general, oil-based fossil fuels are refined prior to their final intended usage. Therein, oil is processed and refined into more useful products, such as petroleum, naphtha, gasoline, diesel fuel, asphalt base, heating oil, kerosene, and liquefied petroleum gas.
5. **Hydrocarbon fuel -** All fossil fuels are hydrocarbons, and certain plant species produce hydrocarbons, which may be separated from the remainder of the plant material.

7.5.2 Physical phases of fuel for combustion

Fuel can take any of three (more or less) phases of matter.

7.5.3 Gaseous fuel

The term 'gaseous fuel' refers to any combustible fuel that exists in the gaseous state under normal temperatures and pressures. Gaseous fuels are typically composed of a wide-range of chemical compounds. Low boiling point hydrocarbons (both paraffins and olefins), hydrogen, carbon monoxide, and inert gases (nitrogen and carbon dioxide) are among the many chemical constituents of common gaseous fuels.

Examples of common gaseous fuels include:

1. **Natural gas (fossil fuel)** is a gaseous fossil fuel that is formed naturally beneath the Earth and is typically found with or near crude oil reservoirs. Natural gas consists of a fluctuating range of low boiling point hydrocarbons. Methane is the primary chemical component, and can be present in amounts ranging from 70% to 99.6% by volume. Ethane can be present in amounts ranging from 2% to 16% by volume. Natural gas consists primarily of methane (CH4). The heat is released as the carbon (C) and hydrogen (H2) combine (react) with oxygen and produce water (H20) and carbon dioxide (CO2).Carbon dioxide, nitrogen, hydrogen, oxygen, propane, butane, and heavier hydrocarbons are also typically present in the fuel analysis.3 The exact analysis usually varies somewhat depending

on the source of the gas and on any heating value adjustments or supplementation."

2. **Liquified petroleum gas (LPG)** is the general term used to describe a hydrocarbon that is stored as a liquid under moderate pressure, but is a gas under normal atmospheric conditions. The primary chemical components of LPG are propane, propylene, normal butane, isobutane, and butylene.
3. **Gaseous biofuel** - Gaseous fuels produced from biomass.

7.5.4 Liquid (and oil) fuels

The term 'liquid fuel' refers to any combustible fuel that exists in the liquid state under normal temperatures and pressures. By definition, a liquid fuel is a fuel substance that deforms continuously when shear stress is applied. The most common liquid fuels are hydrocarbon-based, but there are many non-hydrocarbon-based liquid fuels, such as 100% hydrogen-peroxide.

The combustion of a liquid fuel in an oxidizing atmosphere happens in the gas phase. It is the vapor [of the liquid] that burns, not the liquid. The liquid, however, can still combust so rapidly that it explodes. A liquid fuel will normally catch fire only above a certain temperature: its flash point. The flash point of a liquid fuel is the lowest temperature at which it can form an ignitable mix with air. It is the minimum temperature at which there is enough evaporated fuel in the air to start combustion.

In general, liquid fuels are combusted in devices called 'burners', with a pre-combustion atomization phase using an appropriate fuel atomizer. Atomization is the process of breaking up bulk liquid into many small droplets (i.e., "spraying" the liquid). In order to have good combustion, fuel and air must mix well. A bulk liquid fuel has a limited surface area to contact with the air. This is the reason that liquid fuel, requires atomization before burning. Oils, in particular, must be atomized for optimal combustion.

Examples of liquid fuels include:

1. **Crude oil (fossil fuel)** - The primary chemical components of crude oil are carbon, hydrogen, sulfur, oxygen, and nitrogen. The percentages of these elements found in a crude oil are most frequently used to characterize the oil. Crude oils also contain inorganic elements such as vanadium, nickel, and sodium, and usually contain some amount of water and ash (noncombustible material). The main hydrocarbon constituents of crude oils are alkanes (paraffins), cycloalkanes (naphthenes), and aromatics. The end products derived from crude oil number in the thousands.
2. **Fuel oils (a.k.a., marine fuel or furnace oil)** - Fuel oil is a fraction obtained from petroleum distillation, either as a distillate or a residue. It is the most common fuel on the planet today. If the petroleum context of the term is removed, then a 'fuel oil' could otherwise be defined as any liquid fuel that is burned in a furnace or boiler for the generation of heat or used in an engine for the generation of power, except oils having a flash point of approximately 40°C, and oils burned in cotton or wool-wick burners. In the hydrocarbon context, 'fuel oil' is made of long hydrocarbon chains, particularly alkanes, cycloalkanes and aromatics. The term fuel oil is also used in a stricter sense to refer only to the heaviest commercial fuel that can be obtained from crude oil (i.e., heavier than gasoline and naphtha). The two classifications that separate hydrocarbon fuel oils are "distillates" and "residuals," where distillates indicate a distillation overhead product (lighter oils) and residuals indicate a distillation bottom product (heavier oils).
3. **Liquid biofuel** - Liquid fuels produced from biomass.

NOTE: *Fossil-based oil and natural gas are found under ground between folds of rock and in areas of rock that are porous and contain the oils within the rock itself. The folds of rock were formed as the earth shifts and moves. It's similar to how a small, throw carpet will bunch up in places on the floor.*

7.5.5 Solid fuel

Solid fuels often need to be prepared (e.g., pulverized and grinding) to increase surface area for more efficient combustion. The primary furnace considerations when firing solid fuels revolve around the high levels of ash that are generated. All solid fuels oxidize in a similar manner: the flow of processes are heat-up, devolatilization, volatile oxidation, and finally char burnout. With a solid fuel, the act of combustion consists of three relatively distinct but overlapping phases:

1. <u>Preheating phase</u> - when the unburned fuel is heated up to its flash point and then fire point. Flammable gases start being evolved in a process similar to dry distillation. All solid fuels require initial heat-up as the first step in oxidation to dry the material. Depending on the fuel type, swelling, shrinking, and breakup may occur partially at this stage.
2. <u>Distillation phase or gaseous phase</u> - when the mix of evolved flammable gases with oxygen is ignited. Energy is produced in the form of heat and light. Flames are often visible. Heat transfer from the combustion to the solid maintains the evolution of

flammable vapours.

3. Charcoal phase or solid phase - when the output of flammable gases from the material is too low for persistent presence of flame and the charred fuel does not burn rapidly and just glows and later only smoulders

Examples of solid fuels include:

1. **Biomass** and **solid biofuel** - Raw biomass and solid fuels produced from biomass. Note that raw biomass is technically a fuel source, but it is not a processed or concentrated fuel source.
   - **Pete (turf)** - Peat (turf) is an accumulation of partially decayed vegetation or organic matter (i.e., partially composted organic matter) that is unique to natural areas called peatlands, bogs, or mires. Peat has a high carbon content and can burn under low moisture conditions. Once ignited by the presence of a heat source (e.g., a wildfire penetrating the subsurface), it smolders.
2. **Coal** - A combustible black or brownish-black sedimentary rock usually occurring in rock strata in layers or veins called coal beds or coal seams. The harder forms, such as anthracite coal, can be regarded as metamorphic rock because of later exposure to elevated temperature and pressure. Coal is composed primarily of carbon, along with variable quantities of other elements, chiefly hydrogen, sulfur, oxygen, and nitrogen. A fossil fuel, coal forms when dead plant matter is converted into peat, which in turn is converted into lignite, then sub-bituminous coal, after that bituminous coal, and lastly anthracite. This involves biological and geological processes that take place over time. Coal may be ranked by class from oldest to youngest, geologically. Coal is, relatively speaking, pure carbons. Coal starts initially forming from vegetation and wood under pressure and temperature over a long period of time. During this time period, the initial formation goes from humates to anaerobic and then peat. Final formation to coal then proceeds in order to yield lignite, subbituminous, bituminous and finally anthracotic coal over time. Coal is a hard, black colored rock-like substance. It is made up of carbon, hydrogen, oxygen, nitrogen and varying amounts of sulphur, as well as many other minerals in lesser amounts (including, mercury). There are three main types of coal – anthracite, bituminous and lignite. Anthracite coal is the hardest and has more carbon, which gives it a higher energy content. Lignite is the softest and is low in carbon but high in hydrogen and oxygen content. Bituminous is in between.
3. **Petroleum coke (**abbreviated **pete coke** or **petcoke)** - is a carbonaceous solid delivered from oil refinery coker units or other cracking processes.

7.5.6 Combustion power production systems

> **NOTE:** *Combustion is currently the only currently known power source capable of placing objects in earth's orbit, by powering rockets. These rockets are often powered by liquid ~100% hydrogen peroxide (H_20_2 is a type of rocket fuel). This fuel is atomized before combustion.*

Combustion systems include, but are not limited to the following primary types:

1. **Match system** - One end is of a combustible material (e.g., wood) is coated with another material that can be ignited by frictional heat generated by "striking" the material against a suitable surface.
2. **Wick system** - A wick is usually a braided textile (e.g., hemp or cotton) that holds the flame of a candle or oil lamp for a set period of time depending upon the amount of wick. A candle wick works by capillary action, conveying ("wicking") the fuel to the flame.
3. **Firepit system** - A structure made to contain a fire (outside). In general, firepits are designed for the combustion of biomass in an open space.
4. **Fireplace** or **firebox** - A enclosure/structure made of brick, stone or metal designed to contain a fire. Therein, a chimney or other flue allows exhaust to escape. There are three primary types of fireplaces: biomass combustion; gas combustion; and non-combustion electric fireplaces.
5. **Burner system** - A device/structure responsible for: (1) proper mixing of fuel and air in the correct proportions, for efficient and complete combustion; and (2) determining the shape and direction of the flame. The burner is where combustion takes place; where fuel is combusted with an oxidizer to transfer ("convert") the chemical energy in the fuel into electromagnetic [and thermal] energy. In other words, a burner is the part of the equipment where the fuel is actually burned/combusted; it combusts fuel and generates products of combustion, EM radiation (light), heat, and a change in the surrounding fluid. When the fuel is a liquid or gas, then it flows into the burner and is burnt with the oxidizer (e.g., air), which may be provided by a blower (e.g., air blower). Note that the word 'burner' may also be used in a more general sense to describe the overall apparatus in which the fuel is burned and heat is produced (a.k.a., a furnace). Hence, it may be said

that burners for [utility] boilers are designed to efficiently transfer ("convert") the chemical energy of a fuel into heat within the space provided by the boiler's radiant section, also called the 'furnace' (i.e., a boiler-furnace system). A given combustion system may have a single burner or many burners, depending on the size and type of the application. In concern to a boiler, the 'firing rate' of the burner defines the boiler's output (as steam or other). The burner's design and placement must be engineered to achieve the boiler's essential operation. Burners have a variety of applications, including but not limited to: heating liquid in a boiler; incinerating material in an incinerator; and producing heat for a furnace.

A. **Furnace** or **kiln** - An enclosure/structure in which a fuel (independent of its state of matter) is converted to a high temperature heat.
B. **Incinerator** - An incinerator is a furnace for burning waste. Modern incinerators include pollution mitigation equipment such as flue gas cleaning. There are various types of incinerator plant design: moving grate, fixed grate, rotary-kiln, and fluidised bed.
C. **Boiler** - An enclosed vessel in which water or other fluid is heated (and possibly, circulated) for a separate purpose/function. The fluid does not necessarily boil. The heated or vaporized fluid exits the boiler for use in various processes or heating applications. Hence, there are many sub-types of boiler. A kettle is the most elementary form of a boiler. Common uses for a boiler include, but are not limited to: producing hot water or steam for heating (e.g., water heating, central heating, or cooking); producing steam for use within a manufacturing facility (e.g., atomizing oil for oil-fired burners or for sanitizing equipment); and producing steam to turn a turbine and generate electrical power (i.e., boiler-based power generation). The most common combustible materials (or fuels) used in heating boilers are oil and gas.

NOTE: *Sometimes the word 'boiler' is used in a way that it includes the burner (or heater or furnace) component.*

Combustion systems can be chambered or unchambered. A chamber is an enclosed space specifically for combustion:

- **Chambered system** - A combustion chamber is a space where a fuel/air mixture is burned. Combustion chambers are found in all internal combustion engines (ICEs). Internal combustion engines include, but are not limited to: petrol (gasoline) engines; diesel engines; gas turbines and jet engines (therein, called a combustor); rocket engines. The term combustion chamber is also used to refer to an additional space between the firebox and boiler in a steam locomotive. 'Micro combustion chambers' are the devices in which combustion happens at a very small volume, due to which surface to volume ratio increases which plays a vital role in stabilizing the flame.
- **Combustor (a.k.a., burner)** - A combustor is a component or area of a gas turbine, ramjet, or scramjet engine where combustion takes place. It is also known as a burner, combustion chamber or flame holder.

7.5.7 The ideal combustion system

An "ideal" combustion/fuel burning system would have [at least] the following characteristics:

1. No excess oxygen or unburned combustibles in the end products of combustion.
2. A low rate of auxiliary ignition-energy input to initiate the combustion process.
3. An economic reaction rate between fuel and oxygen compatible with acceptable nitrogen and sulfur oxide formation.
4. An effective method of handling and disposing of the solid impurities introduced with the fuel.
5. Uniform distribution of the product weight and temperature in relation to the parallel circuits of heat absorbing surface.
6. A wide and stable firing range, fast response to changes in firing rate, and high equipment availability with low maintenance.

7.6 Environmental impact

NOTE: *Combustion can be used to destroy (incinerate) waste, both non-hazardous and hazardous.*

Combustion/burning/fire and its by-products can be extremely dangerous to living organisms and ecosystems. All forms of combustion must be carefully monitored and unintentional/undesirable combustive processes must be safely extinguished. In a human habitat, fire control/"fighting" services must exist to extinguish and/or contain uncontrolled fires. Further, fire prevention as a design consideration, is intended to reduce sources and/or the probability of ignition. Fire prevention also includes education to facilitate knowledge of what causes fires, how to avoid fires, and what to do if an uncontrolled fire occurs.

Uncontrolled fires are extremely dangerous to living organisms and other environmental materials. Burning releases carbon monoxide, carbon dioxide, nitrogen oxides, and other pollutants and particulates. If these pollutants are not captured and recycled, burning can create smog. Emission from combustion will enter the atmosphere and spread for hundreds, if not thousands of kilometres. Coal, for example, is often high in mercury. Emissions from coal-fired systems account for 13 to 26 percent of the total (natural plus anthropogenic) airborne emissions of mercury in various locals.

8 Hydropower (water power)

> **DEFINITION: Hydrokinetic** *technologies and devices produce power by harnessing the kinetic energy of a body of water (i.e., the energy contained in its motion).*

Hydropower or water power (from the Greek: ύδωρ, "water") is power derived from the energy of falling or otherwise fast moving water, which may be harnessed for useful purposes. Hydropower refers to the power that is produced by the pressure of moving water (i.e., force of water moving at a velocity). However, it should be noted here that although moving water has kinetic energy, water itself is a carrier of energy. Note here that hydropower is similar to hydraulic power, except that hydraulic power more generally refers to any intentionally pressurized liquid (there is some overlap between hydro- and hydraulic-power). Hydropower has been used since ancient times to grind flour, irrigate, and perform other tasks (e.g., watermills). Through hydropower technology, the potential and/or kinetic energy of water is transferred (i.e., harnessed) to mechanical power, and then, to useful work or to electrical (or other) energy. When hydropower is used to produce electric power it is known as hydroelectric power (hydroelectricity).

> **HISTORICALLY NOTE:** *Hydropower has been used for thousands of years to mill grain. In a 'tide mill', the incoming tidal water is contained in large storage ponds, and as tide goes out (recedes), its movement/pressure turns a waterwheels that uses the mechanical power to mill grain.*

Hydro power/energy is available in many forms: potential energy from high "heads" of water (i.e., water at elevation) retained in dams; kinetic energy from current flow in rivers and tidal barrages; and kinetic energy also from the movement of waves on relatively static water masses. Most ways of harnessing this energy involve directing the water flow through a turbine to generate mechanical power, which may then be used to generate electric power. Those hydropower harnessing systems that do not use a turbine, usually involve using the movement of the water to drive some other form of hydraulic or pneumatic mechanism to perform the same task.

> **TERMINOLOGY:** *The difference in height between the water source and the water outflow is known as 'head', and the potential energy of the water is directly proportional to the 'head'.*

Generally speaking, every form of hydropower is originally derived from one or more of the following other sources of energy: solar, geothermal, and/or "gravitational"/planetary. Specifically, when hydropower takes the form of water moving over or through land (i.e., "running" water), then it may be considered a form of solar energy, as the sun drives water evaporation from the ocean, and winds carry the moisture overland. Similarly, when hydropower takes the form of waves, then it may be considered a form of solar energy since the wind is the most significant factor in wave generation, and the wind comes from the interface of solar energy with the Earth. When hydropower takes the form of an ocean current or tide, then it may be considered a form of "gravitational" energy, due to its movement being significantly derived from the pull of the Moon (and Sun) on water. When hydropower comes from thermal differentials in water, then it may be considered a form of solar and/or geothermal energy, since temperature differentials in the Earth, and sunlight, generate the movement of the water.

NOTE: *Planetary energy refers to the interactive force between the earth, moon, and sun, causing a periodical state change in natural water reservoirs called 'tide'.*

8.7 Hydro-electric power

Hydro-electric power (hydroelectric or hydroelectricity) is a form of hydropower. Hydroelectric power is generated by harnessing/controlling the power of moving water (water pressure, mechanical energy) to produce electric power. In most cases, hydroelectric power (hydroelectricity) is generated through the mechanically powered rotation of a turbine connected via shaft [power] to a generator that produces electric power through the electromagnetic induction effect.

NOTE: *Electrolysis and electrodialysis involve the production of electrical power through the use of water, they are not technically hydropower sources, because they do not produce electric current through the direct movement/pressure of water.*

Hydroelectric power may be sub-classified based on the location of the water used and the direct source of the water's movement. There are two top level location-based categories of hydroelectric power:

- Land
- Ocean/marine

In general, ocean energy can be sub-classified into six types of different origin and characteristics:

- Ocean wave (wave)
- Tidal range (tidal)
- Tidal current (current)
- Ocean current (current)
- Ocean thermal energy (thermal)
- Salinity gradient (osmotic)

A [water] 'current' is a relatively large movement of water in one direction. In a current, water is moving forward. Water passing over/through land, and water in the ocean, can have a current. In the ocean, currents can be temporary or long-lasting; they can be near the surface or in the deep ocean. Ocean currents are driven by several factors, including gravitational/planetary motion and thermohaline circulation. Thermohaline circulation generates large ocean currents driven by differences in temperature (thermo) and salinity (haline). A current, in a river or stream, is the flow of water influenced by gravity as the water moves downhill to reduce its potential energy. The current varies spatially as well as temporally within the stream, dependent upon the flow volume of water, stream gradient, and channel geometrics. The term 'current' (as in, 'water current') can be applied in three ways to hydroelectric power:

1. Current as tide (tidal current)
2. Current as stream/river (river/stream current)
3. Controlled current as dam

The form (i.e., "sculpting") of water into waves (located on the surface of a body of water) is commonly caused by wind transferring its energy to the water. Large surface waves, known as 'swells', can travel over long distances. A surface wave's size depends on wind speed, wind duration, and the area over which the wind is blowing (the 'fetch'). Tides may be viewed as waves; the largest waves on the planet, and they cause the sea to rise and fall along the shore around the world. Tides exist due to the gravitational/planetary pull of the moon and the sun, but vary depending on where the moon and sun are in relation to the ocean as the earth rotates on its axis.

8.7.1 Hydroelectric power generation

Hydroelectric power generation systems can be classified in the following ways:

1. According to the availability of head (elevation drop of water):
   A. High head power
   B. Medium head power
   C. Low head power
2. According to the nature of load:
   A. Base load generation
   B. Peak load generation
3. According to capacity (quantity of water available):
   A. Large (>100MW); medium (25-100MW); small (1-25MW); mini (100KW-1MW); micro (5-100KW); pico (<5KW)
4. According to hydrological region:
   A. Single
   B. Cascade
5. According to transmission system:
   A. Isolated
   B. Connected to grid
6. As land - according to quantity of water available:
   A. Hydroelectric generation with storage reservoirs

(controlled current as dam)
   B. Run of river generation without pondage (current as river/stream)
   C. Run of river generation with pondage (current as river/stream)
   D. Pump storage
7. As ocean/marine:
   A. Current (current in general)
   B. Tidal (current as tide)
   C. Wave
   D. Thermal
   E. Osmotic

8.7.2 Land-based hydroelectric power

There are three basic landed ways in which hydroelectric power may be generated from water. They all involve the use of a turbine-electric system. Land-based hydroelectric power generation systems can be classified according to the characteristics of the watercourse which is being used as a power (energy) source:

1. **Dammed-hydro** (hydroelectric dams; reservoir-type; impoundment power station) - The potential energy of water is collected in a dam. A dam is used to store river water in a reservoir. Water is released from the dam in a controlled manner. Water released from the reservoir flows/falls through a turbine, spinning it, which in turn powers an electric generator to produce electricity. A hydroelectric dam installation uses the potential energy of the water retained in the dam, and its release as kinetic energy, to drive a water turbine, which in turn drives an electric generator. The available energy therefore depends on the 'head' of the water (i.e., elevation) above the turbine and the volume of water flowing through it. Turbines used for this purpose are usually reaction type, whose blades are fully submerged in the water flow. The height of the dam and mass of water behind the dam (as well as the turbine-electric system) determines power output. Herein, available power is expressed as:
   - Potential energy per unit volume = ρgh
   - Where, ρ is he density of the water (103 Kg/m3), h is the head of water and g is the gravitational constant (10 m/sec2)
   - Where, Q is the volume of water flowing per second (the flow rate in m3/second) and η is the efficiency of the turbine.
   - The power P from a dam is given by: $P = \eta\rho g h Q$
2. **Run-Of-The-River** (ROR; in-stream; diversion-type; channel-type) - A diversion in a river/stream channels/diverts a portion of a water through a canal or penstock to a turbine. The kinetic energy of the flowing water is used to drive the turbine. It generally does not require the use of a dam. An ROR system may or may not have pondage associated with it. 'Pondage' usually refers to a relatively small water storage area behind the weir of a run-of-the-river hydroelectric power plant. A 'weir' is a barrier across a river designed to alter its flow characteristics. Therein, the 'head' is often zero (or close to zero). The available energy therefore depends on the quantity of water flowing through the turbine and the square of its velocity. Impulse turbines, which are only partially submerged, are more commonly employed in fast flowing run of river installations. In deeper, slower flowing rivers, submerged Kaplan turbines may be used to extract the energy from the water flow. Herein, available power is expressed as:
   - The maximum power output from a turbine used in a run of river application is equal to the kinetic energy (½mv^2) of the water impinging on the blades. Taking the efficiency η of the turbine and its installation into account, the maximum output power P_{max} is given by: $P_{max} = ½\eta\rho Q v^2$
   - Where v is the velocity of the water flow and Q is the volume of water flowing through the turbine per second.
   - Q is given by: Q = A v
   - Where A is the swept area of the turbine blades.
   - Thus, $P_{max} = ½\eta\rho A v^3$
   - Note that the power output is proportional to the cube of the velocity of the water.
3. **Pumped storage** - Under this method, electric current is generated through a turbine-generator by intentionally moving water between reservoirs located at different heights. This method is useful for supplying electricity on occasions of high peak demands. When the demand is high, water is released from a higher to a lower reservoir, and run through a turbine-electric system to generate electrical power. During times of low demand, water from lower reservoirs is pumped (using electric power and fluid pressure) up into higher reservoirs. Pumped storage works like a battery, storing potential energy "in" water until it is needed. A pumped storage system can be independent of other land-based hydroelectric production systems, or connected to these systems.

A land-based hydro-electric turbine uses water pressure to turn a generator before the water flows out at very low pressure through a 'tail race'. There is no heat involved. This is not a heat engine. The following turbines are most commonly used in land-based hydroelectric power systems:

1. Francis turbines - suitable for middle-sized available heads. An inward-flow reaction turbine that combines radial and axial flow concepts.
   - Kaplan turbines - suitable for low available heads and larger water flows.
2. Pelton turbines - suitable for high available heads and smaller water flows. A type of impulse turbine

Each type of land-based hydroelectric plant has advantages and disadvantages. It is important to recognize that whenever there is a turbine in a natural water source, then the turbine can/will pose a danger to aquatic organisms.

Land-based hydroelectric systems, dams in particular, have several advantages, including:

1. Technically, "non-pollutive"
2. Climatically renewable
3. Possibility of use as flood control
4. Multiple crop cultivations per year
5. New ecosystem

Land-based hydroelectric systems, dams in particular, have several disadvantages, including:

1. May require the construction of a dam/reservoir, which will modify the local habitat.
2. If a drought occurs, the power station may not have an energy source.
3. Turbines can pose a danger to aquatic lifeforms, and human swimming, and dams pose an obstacle to the movement of aquatic life.
4. Dams have other drawbacks, including:
   A. Loss of nutrient flow down river.
   B. Loss of sediment flow down river.
   C. Sedimentation behind the dam limits lifetime of the dam and increases maintenance requirements.
   D. Flooding of scenic areas and alteration of ecology to create the dam system.
   E. Ecosystem below the dam is usually changed.
   F. Colder, nutrient poor water.
   G. Aesthetics and ecological change as the loss of wild rivers.

8.7.3 Ocean/marine-based hydroelectric: Tidal power

Tidal power, also called "tidal energy", is a form of hydropower that converts the energy obtained from tides into useful forms of power, mainly electric power. Tidal power uses the energy available from the ocean's tidal motions. Tidal motions represent the cyclical rise and fall of water due to tidal phenomena. Tides are highly predictable, and significantly more predictable than wind power and solar power.

In general, tides cycle every 12.5 hours, so daily peak production times and slack times vary, this causes a mismatch in supply and demand, as industrial demand is high during the day and low at night. Tides vary seasonally and monthly as well. The general 'tidal range' around the earth is about 2 feet to about 20 feet, the higher the 'tide range', the more useful the energy.

A tidal stream generator, often referred to as a tidal energy converter (TEC), is a general term for a machine that extracts energy from moving masses of water, in particular tides, although the term is often used in reference to machines designed to extract energy from run-of-river and tidal estuarine sites. Certain types of these machines function very much like underwater wind turbines, and are thus often referred to as tidal turbines.

Tidal power can be classified into four generating methods:

1. **Tidal stream generator (TSG)** - use the kinetic energy of moving water to power turbines. Tidal stream generators can be built into the structures of existing bridges, entirely or partially submersed, thus avoiding concerns over impact on the natural landscape. Land constrictions such as straits or inlets can create high velocities at specific sites, which can be captured with the use of turbines. These turbines can be horizontal, vertical, open, or ducted, and are typically placed near the bottom of the water column (entirely submerged) where tidal velocities are greatest.
2. **Tidal barrage** - use the potential energy in the difference in height (or hydraulic head) between high and low tides. When using tidal barrages to generate power, the potential energy from a tide is harvested through strategic placement of specialized dams. When the sea level rises and the tide begins to come in, the temporary increase in tidal power is channelled into a large basin behind the dam, holding a large amount of potential energy. With the receding tide, this energy is then transferred into mechanical power as the water is released through turbines. Barrages are essentially dams across the full width of a tidal estuary.
3. **Dynamic tidal power (DTP)** - uses an interaction between potential and kinetic energies in tidal flows. It involves the construction of sea dams (30–50 km length) from coasts straight out into the sea or ocean, without enclosing an area. Tidal phase differences are introduced across the dam, leading to a significant water-level differential in shallow coastal seas.
4. **Tidal lagoon** - generally, uses constructed circular retaining walls embedded with turbines that can capture the potential energy of tides. The created

reservoirs are similar to those of tidal barrages, except that the location is artificial and does not contain a pre-existing ecosystem. The lagoons can also be in double (or triple) format without pumping or with pumping that will flatten out the power output. The pumping power could be provided by excess to grid demand renewable energy from for example wind turbines or solar photovoltaic arrays. Excess energy rather than being curtailed could be used and stored for a later period of time. Geographically dispersed tidal lagoons with a time delay between peak production would also flatten out peak production providing near base load production.

The European Marine Energy Centre recognizes six principal types of tidal energy converter (TEC). The types/characteristics of tidal turbines are as follows (Note: any given turbine can have more than one of these characteristics):

1. **Axial turbines** - These are close in concept to traditional windmills, but operating under the sea.
   A. **Horizontal axis turbines** - the main rotor shaft (and generally, electrical generator) are pointed into the current flow (generally set horizontally, but not necessarily horizontally).
   B. **Vertical axis turbines** - the main rotor shaft is set transverse to the flow (generally set vertically, but not necessarily vertically). In other words, the axis is positioned perpendicular to current flow.
2. **Crossflow turbines** (Banki-Michell turbine, Ossberger turbine) - These are impulse turbines; water flows through the runner transversely, striking the blades once on entry and meeting them again as it leaves the runner. Unlike most water turbines, which have axial or radial flows, in a crossflow turbine the water passes through the turbine transversely, or across the turbine blades. As with a waterwheel, the water is admitted at the turbine's edge. After passing the runner, it leaves on the opposite side. Going through the runner twice provides additional efficiency. When the water leaves the runner, it also helps clean the runner of small debris and pollution. The cross-flow turbine is a low-speed machine. These turbines can be deployed either vertically or horizontally. Crossflow turbines are often constructed as two turbines of different capacity that share the same shaft. The turbine wheels are the same diameter, but different lengths to handle different volumes at the same pressure.
3. **Flow augmented turbines** - Turbine that use flow augmentation measures (e.g., a duct or shroud) such that the incident power available to the turbine can be increased. The most common example uses a shroud to increase the flow rate through the turbine, which can be either axial or crossflow.
4. **Shrouded tidal turbine** - a turbine enclosed in a venturi shaped shroud or duct (ventiduct), producing a sub atmosphere of low pressure behind the turbine. The venturi shrouded turbine is not subject to the Betz limit and allows the turbine to operate at higher efficiencies than the turbine alone by increasing the volume of the flow over the turbine. The performance of a shrouded turbine varies with the design of the shroud. The available power from a shrouded tidal turbine is expressed as:
   - The maximum power output from a shrouded water turbine used in tidal energy applications is equal to the kinetic energy of the water impinging on the blades, similar to the "run of river" calculation. Taking the efficiency η of the turbine and its installation into account, the maximum output power P_{max} is given by
   - $P_{max} = ½ηρAv3$
   - Where, v is the velocity of the water flow and A is the swept area of the blades.
5. **Oscillating hydrofoils** - Oscillating devices do not have a rotating component, instead they use aerofoil sections that oscillate (pushed sideways) by the flow of water. In other words, oscillating hydrofoils are a form of hydroelectric generation system in that they are not turbines. However, they still use the electromagnetic induction effect to generate electric power. Oscillating stream power extraction was proven with the omni- or bi-directional winged pump windmill.
6. **Venturi devices** - These devices use a shroud or duct in order to generate a pressure differential which is used to run a secondary hydraulic circuit which is used to generate power. These devices make use of the venturi effect.
7. **Archimedes screws** - Water is pumped by turning a screw-shaped surface inside a pipe. A machine historically used for transferring water from a low-lying body of water into irrigation ditches.
8. **Tidal kites** - A tidal kite turbine is an underwater kite system or paravane that converts tidal energy into electricity by moving through the tidal stream. The kite is tethered by a cable to a fixed point. It "flies" through the current carrying a turbine. It moves in a figure-eight loop to increase the speed of the water flowing through the turbine tenfold.

These systems have several possible mounting positions:

1. **Bottom-mounted** - mounted stationary to the sea floor/bed.
2. **Fully underwater cable tethered** - tethered via cables connected to the sea floor, while the turbine remains fully in the water.
3. **Surface floating cable tethered** - the turbine is connected to a floating platform, which is tethered to the sea floor.
4. **Structure mounted** - mounted onto an architectural structure, such as a bridge.

8.7.4 Ocean/marine-based hydroelectric: Wave power

Wave power uses the energy available from the ocean's surface wave motion to generate electrical power. There are several types of devices design to harvest wave power for the generation of electric power:

1. **Oscillating float system** - a float is housed inside an cylinder shaped buoy which is open at the bottom, and moored to the seabed. Inside the cylinder the float moves up and down on the surface of the waves as they pass through the buoy. Various methods have been employed to turn the motion of the float into electrical energy. These include:
   A. Hydraulic systems in which air is compressed in a pneumatic reservoir above the float during its upward movement on the crests of the waves. After the crests have passed, the air expands and forces the float downwards into the following troughs of the waves. A hydraulic system then uses the reciprocating movement of the float to pump water through a water turbine which drives a rotary electrical generator. Instead of generating the electricity on board the buoy, some systems pump the hydraulic fluid ashore to power shore based generators.
   B. Pneumatic systems in which the air displaced in the cylinder is used to power an air turbine which drives the generator.
   C. Linear generators to turn the reciprocating motion of the float directly into electrical power.
2. **Oscillating paddle system** - uses large paddles moored to the ocean floor to mimic the swaying motion of sea plants in the presence of ocean waves. The paddles are fixed to special hinged joints at the base which use the swaying motion of the paddles to pump water through a turbine generator.
3. **Oscillating snake system** - uses a series of floating cylindrical sections linked by hinged joints. The floating snake is tethered to the sea bed and maintains a position head on into the waves. The wave-induced motion at the hinges is used to pump high-pressure oil through hydraulic motors via smoothing accumulators. The hydraulic motors in turn drive electrical generators to produce the electrical power.
4. **Oscillating water column** - waves enter and exit a partially submerged collector from below, causing the water column inside the collector to rise and fall. The changing water level acts like a piston as it drives air that is trapped in the device above the water into a turbine, producing electricity via a coupled generator. Water columns are often formed within large concrete structures built on the shore line, or on rafts. The structure is open at both the top and the bottom. The lower end is submerged in the sea and an air turbine fills the aperture at the top. The rising and falling of the water column inside the structure moves the air column above it driving the air through the turbine generator. The turbine has movable vanes which rotate to maintain unidirectional rotation when the movement of the air column reverses.
5. **Pressure transducer system** - uses a submerged gas-filled tank with rigid sides and base and a flexible, bellows-like, top. The gas in the tank compresses and expands in response to pressure changes from the waves passing overhead causing the top to rise and fall. A lever attached to centre of the top drives pistons, which pump pressurized water ashore for driving hydraulic generators.
6. **Wave capture system** - use a narrowing ramp to funnel waves into an elevated reservoir. Waves entering the funnel over a wide front are concentrated into a narrowing channel which causes the amplitude of the wave to increase. The increased wave height coupled with the momentum of the water is sufficient to raise a quantity of water up a ramp and into a reservoir situated above the sea level. Water form the reservoir can then be released through a hydroelectric turbine located below the reservoir to generate electric power.
7. **Overtopping wave system** - channel waves onto a tapered ramp that causes an increase in their amplitude. The crests of the waves overtop the ramp and spill into a low dam. Water from the low dam then flows through hydroelectric turbines back into the sea beneath the floating structure. A floating reservoir, in effect, is formed as waves break over the walls of the device. The reservoir creates a head of water—a water level higher than that of the surrounding ocean surface—which generates the pressure necessary to turn a hydro

turbine as the water flows out the bottom of the device, back into the sea.

8. **Lever system** - long levers may be mounted on steel piles or on floating platforms. Large floats or buoys are attached to the extremities of the levers which move up and down with the waves. The movement of the lever arms forces fluid into a central hydraulic accumulator and through to a generator turbine. Alternatively high-pressure water can be pumped ashore to power shore based generators.
9. **Point absorber system** - utilizes wave energy from all directions at a single point by using the vertical motion of waves to act as a pump that pressurizes seawater or an internal fluid, which drives a turbine. This type of device has many possible configurations. One configuration, called a hose pump point absorber, consists of a surface-floating buoy anchored to the sea floor, with the turbine device as part of the vertical connection. The wave-induced vertical motion of the buoy causes the connection to expand and contract, producing the necessary pumping action. Through engineering to generate device-wave resonance, energy capture and electricity generation by point absorbers can be maximized.
10. **Attenuator system (heave-surge devices)** - are long, jointed floating structures are aligned parallel to the wave direction and generate electricity by riding the waves. The device, anchored at each end, utilizes passing waves to set each section into rotational motion relative to the next segment. Their relative motion, concentrated at the joints between the segments, is used to pressurize a hydraulic piston that drives fluids through a motor, which turns the coupled generator.

The available power in a wave powered electrical generation system is expressed as:

1. The wave power per unit length of the wave front P_L is given by: $P_L = \rho g a^2 \lambda / 4T$
2. where, ρ is he density of the water (e.g., 103 Kg/m3), a is the wave amplitude (half of the wave height), g is the gravitational constant (10 m/sec^2), and λ is the wave length of the oscillation and T the period of the wave.

8.7.5 Ocean/marine-based hydroelectric: Osmotic power

Osmotic power (salinity gradient power or "blue energy") is the energy available from the difference in the salt concentration between seawater and river water. There are several methods for generating electric power from the salinity gradient. The key waste product is brackish water. This byproduct is the result of natural forces that are being harnessed: the flow of fresh water into seas that are made up of salt water.

1. **Reversed electrodialysis (RED)** - A process that relies on osmosis with ion specific membranes.
2. **Pressure retarded osmosis (PRO)** - A process that relies on osmosis with ion specific membranes. Seawater is pumped into a pressure chamber where the pressure is lower than the difference between fresh and salt water pressure. Fresh water moves in a semipermeable membrane and increases its volume in the chamber. As the pressure in the chamber is compensated a turbine spins to generate electricity.
3. **The capacitive method** - With this method energy can be extracted out of the mixing of saline water and freshwater by cyclically charging up electrodes in contact with saline water, followed by a discharge in freshwater. Each completed cycle effectively produces energy.
4. **Vapor pressure differences** - Does not rely on membranes, so filtration requirements are not as important as they are in the PRO and RED methods.
   A. Open cycle - Similar to the open cycle in ocean thermal energy conversion (OTEC).
   B. Absorption refrigeration cycle (closed cycle) - For the purpose of dehumidifying air, in a water-spray absorption refrigeration system, water vapor is dissolved into a deliquescent salt water mixture using osmotic power as an intermediary. The primary power source originates from a thermal difference, as part of a thermodynamic heat engine cycle.
5. **Solar pond** - This method does not harness osmotic power, only solar power. Sunlight reaching the bottom of the saltwater pond is absorbed as heat. The effect of natural convection, wherein "heat rises", is blocked using density differences between the three layers that make up the pond, in order to trap heat. The upper convection zone is the uppermost zone, followed by the stable gradient zone, then the bottom thermal zone. The stable gradient zone is the most important. The saltwater in this layer can not rise to the higher zone because the saltwater above has lower salinity and is therefore less-dense and more buoyant; and it can not sink to the lower level because that saltwater is denser. This middle zone, the stable gradient zone, effectively becomes an "insulator" for the bottom layer (although the main purpose is to block natural convection, since water is a poor insulator). This water from the lower layer,

the storage zone, is pumped out and the heat is used to produce energy, usually by turbine in an organic Rankine cycle. A technology called salinity gradient solar pond (SGSP) may be used. In theory a solar pond could be used to generate osmotic power if evaporation from solar heat is used to create a salinity gradient, and the potential energy in this salinity gradient is harnessed directly using one of the first three methods above, such as the capacitive method.

6. **Boron nitride nanotubes** - An impermeable and electrically insulating membrane is pierced by a single boron nitride nanotube with an external diameter of a few dozen nanometers. With this membrane separating a salt water reservoir and a fresh water reservoir, an electric current passes through the membrane using two electrodes immersed in the fluid either side of the nanotube.

8.7.6 Ocean/marine-based hydroelectric: Ocean thermal Energy conversion

Ocean thermal energy conversion (OTEC) is a hydrothermal process that can produce electric power by using the temperature difference between cooler deep water and warmer shallow (or surface seawaters) to run a binary cycle electric generating heat engine. In a heat engine, thermal energy does the work. An OTEC system pumps large quantities of deep cold seawater and surface seawater to run a power cycle and produce electricity. The thermal energy of the warmer oceans of the world can be used to generate electricity in much the same way as geothermal heat is used for electrical energy generation. Warmer water is taken from the surface of the ocean to vaporise the fluid in the turbine circuit. Cold water is pumped from the depths of the ocean to condense the working fluid. OTEC is a base load electricity generation system. In the oceans the temperature difference between surface and deep water is greatest in the tropics, although still a modest 20°C to 25°C. It is therefore in the tropics that OTEC offers the greatest possibilities. OTEC energy harvesting is similar to geothermal energy extraction described above except that the temperature gradient has an opposite slope.

Sea water is heated by energy both from the Sun and from the Earth below. The solar energy falling on the water surface is greater than the heat flow emanating from the Earth so that the temperature at the surface is greater than the temperature in the depths of the water.

In a dual cycle “binary plants” the hot water circuit passing through the thermal source is separated from the closed loop working fluid circuit used in the turbine by a heat exchanger. The hot water gives up its heat in the heat exchanger to a working fluid with a low boiling point and high vapour pressure at low temperatures when compared to steam. The working fluid is typically an organic compound (e.g., ammonia, butane, pentane or isopentane) which circulates through the secondary side of the heat exchanger where it vaporises and the vapour is then used to rotate a turbine in a conventional Rankine cycle electricity generating plant. After the vaporised binary liquid has passed through, and given up its energy to, the turbine it is condensed and recycled for re-use through the heat exchanger.

There are a variety of potential working fluids. Ammonia, which has superior transport properties, is easy availability, but it is toxic and flammable. Fluorinated carbons such as CFCs and HCFCs are not toxic or flammable, but they contribute to ozone layer depletion. Hydrocarbons too are good candidates, but they are highly flammable; in addition, this would create competition for use of them directly as fuels. The power plant size is dependent upon the vapor pressure of the working fluid. With increasing vapor pressure, the size of the turbine and heat exchangers decreases while the wall thickness of the pipe and heat exchangers increase to endure high pressure especially on the evaporator side.

Cold seawater is an integral part of each of the three types of OTEC systems: closed-cycle, open-cycle, and hybrid. To operate, the cold seawater must be brought to the surface. The primary approaches are active pumping and desalination. Desalinating seawater near the sea floor lowers its density, which causes it to rise to the surface.

OTEC systems may be either closed-cycle or open-cycle:

1. **Closed-cycle OTEC (Anderson cycle)** - Uses working fluids with a low boiling point that are typically thought of as refrigerants (e.g., ammonia or R-134a) to power a turbine, which powers a generator. Warm surface seawater is pumped through a heat exchanger to vaporize the fluid. The expanding vapor turns the turbo-generator. Cold water, pumped through a second heat exchanger, condenses the vapor into a liquid, which is then recycled through the system. The most commonly used heat cycle for OTEC to date is the Rankine cycle, using a low-pressure turbine.
2. **Open-cycle OTEC (Claude cycle)** - Uses vapour from the seawater itself as the working fluid -- warm surface water at around 27 °C (81 °F) enters an evaporator at pressure slightly below the saturation pressures causing it to vaporize. Warm seawater is first pumped into a low-pressure container, which causes it to boil. In some schemes, the expanding vapour drives a low-pressure turbine attached to an electrical generator. The vapour, which has left its salt and other contaminants in the low-pressure container, is pure fresh water. It is condensed into a liquid by exposure to cold temperatures from deep-ocean water. This method produces desalinized (desalinated) fresh water, suitable for drinking water, irrigation or

aquaculture. In other schemes, the rising vapour is used in a gas lift technique of lifting water to significant heights. Depending on the embodiment, such vapour lift pump techniques generate power from a hydroelectric turbine either before or after the pump is used.

3. **Hybrid OTEC** - A hybrid cycle combines the features of the closed- and open-cycle systems. In a hybrid, warm seawater enters a vacuum chamber and is flash-evaporated, similar to the open-cycle evaporation process. The steam vaporizes the ammonia working fluid of a closed-cycle loop on the other side of an ammonia vaporizer. The vaporized fluid then drives a turbine. The steam condenses within the heat exchanger and provides desalinated water.

NOTE: *OTEC and electrolysis technologies can be combined to produce electrical power and hydrogen.*

There are three possible locations for the placement of an OTEC plant:

1. **Land-based** and **near-shore** - Systems constructed on or near land do not require sophisticated mooring, lengthy power cables, or the more extensive maintenance associated with open-ocean environments. They can be installed in sheltered areas so that they are relatively safe from storms and heavy seas. Electricity, desalinated water, and cold, nutrient-rich seawater could be transmitted from near-shore facilities via conduits. In addition, land-based or near-shore sites allow plants to operate with related industries such as mariculture or those that require desalinated water. Land-based or near-shore sites can also support mariculture or chilled water agriculture. Tanks or lagoons built on shore allow workers to monitor and control miniature marine environments.
2. **Shelf-based** - To avoid the turbulent surf zone as well as to move closer to the cold-water resource, OTEC plants can be mounted to the continental shelf at depths up to 100 meters (330 ft). A shelf-mounted plant could be towed to the site and affixed to the sea bottom. The complexities of operating an OTEC plant in deeper water may make them more expensive than land-based approaches. Problems include the stress of open-ocean conditions and more difficult product delivery. Addressing strong ocean currents and large waves adds engineering and construction expense. Platforms require extensive pilings to maintain a stable base. Power delivery can require long underwater cables to reach land.

- **Floating [off shore]** - The difficulty of mooring plants in very deep water complicates power delivery. Cables attached to floating platforms are more susceptible to damage, especially during storms. Cables at depths greater than 1000 meters are difficult to maintain and repair. Riser cables, which connect the sea bed and the plant, need to be constructed to resist entanglement. As with shelf-mounted plants, floating plants need a stable base for continuous operation. Major storms and heavy seas can break the vertically suspended cold-water pipe and interrupt warm water intake as well. To help prevent these problems, pipes can be made of flexible polyethylene attached to the bottom of the platform and gimballed with joints or collars. Pipes may need to be uncoupled from the plant to prevent storm damage. As an alternative to a warm-water pipe, surface water can be drawn directly into the platform; however, it is necessary to prevent the intake flow from being damaged or interrupted during violent motions caused by heavy seas.

Technical difficulties include, but are not limited to:

1. **Dissolved gases** - As cold water rises in the intake pipe, the pressure decreases to the point where gas begins to evolve. If a significant amount of gas comes out of solution, placing a gas trap before the direct contact heat exchangers may be justified.
2. **Microbial fouling** - Because raw seawater must pass through the heat exchanger, care must be taken to maintain good thermal conductivity. Biofouling layers as thin as 25 to 50 micrometres (0.00098 to 0.00197 in) can degrade heat exchanger performance by as much as 50%. (Berger, 1986)
3. **Sealing** - The evaporator, turbine, and condenser operate in partial vacuum ranging from 3% to 1% of atmospheric pressure. The system must be carefully sealed to prevent in-leakage of atmospheric air that can degrade or shut down operation. In closed-cycle OTEC, the specific volume of low-pressure steam is very large compared to that of the pressurized working fluid. Components must have large flow areas to ensure steam velocities do not attain excessively high values.
4. **Parasitic power** consumption by exhaust compressor.

8.8 Hydroelectric issues

There are a variety of technical challenges/issues associated with hydroelectric generation, including but not limited to:

1. Ecological/environmental concerns - Tidal power systems can have effects on marine life and marine ecology. Turbines can accidentally kill swimming sea life with their rotating blades, although it is possible to create safety mechanisms that turn off the turbine when marine animals approach. Some marine life may no longer utilize the area if threatened with a constant rotating or noise-making object. The Tethys database provides access to scientific literature and general information on the potential environmental effects of tidal energy. These system can interfere with the migrating fish species.
    A. Tidal turbines - High speed water increases the risk of organisms being pushed near or through these devices leading to entanglement and blade strikes. There is also a concern about how the creation of EMF and acoustic outputs may affect marine organisms. It should be noted that because these devices are in the water, the acoustic output can be greater than those created with offshore wind energy. Depending on the frequency and amplitude of sound generated by the tidal energy devices, this acoustic output can have varying effects on marine mammals (particularly those who echolocate to communicate and navigate in the marine environment, such as dolphins and whales). Tidal energy removal can also cause environmental concerns such as disrupting sediment processes and degrading far-field water quality.
    B. Tidal barrage - Installing a barrage may change the shoreline within the bay or estuary, affecting a large ecosystem that depends on tidal flats. Inhibiting the flow of water in and out of the bay, there may also be less flushing of the bay or estuary, causing additional turbidity (suspended solids) and less saltwater, which may result in the death of fish that act as a vital food source to birds and mammals. Migrating fish may also be unable to access breeding streams, and may attempt to pass through the turbines. There are also acoustic concerns. Shipping accessibility is also a concern.
    C. Tidal lagoon - The main concerns are blade strike on fish attempting to enter the lagoon, acoustic output from turbines, and changes in sedimentation processes. However, all these effects are localized and do not affect the entire estuary or bay.
2. Variability of the sea conditions - Sea conditions are highly variable and the system must be able to cope with a wide-range of wave amplitudes and frequencies as well as changes in the directions of currents.
3. Matching the generating equipment to the wave/current characteristics - Mechanisms are required to convert the power of the irregular oscillating mechanical forces induced by the waves into electrical power (synchronised with the grid). This could involve some power electronics. Hydraulic accumulators can be used in-situ, or on shore, to smooth out the energy delivery to the generator.
4. Housing and mooring the equipment - Substantial housings must be provided to protect the generating equipment from the harsh environment. Holding the installation in place is also particularly difficult in deep water.
5. Corrosion - Materials must either have anti-corrosive properties or be protected from corrosion by salt water and atmospheric salts.
6. Fouling - The biological events that happen when placing any structure in an area of high tidal currents and high biological productivity in the ocean will ensure that the structure becomes an ideal substrate for the growth of marine organisms.
7. Energy transmission - Low loss armoured and insulated cables or high pressure pipes must be developed for delivering the electrical or hydraulic energy back to the shore.
8. Resistance to storm damage - The frequency of occurrence of waves of any particular amplitude follows a Rayleigh distribution similar to that which applies to wind speeds. Though the frequency of serious storms may be rather small, a wave of ten times the average amplitude may be expected once every 50 years. From the power calculation below, the wave power is proportional to the square of the wave amplitude. This means that the installation must be designed to withstand forces one hundred times greater than the normal working level.
9. Ship traffic course correction - Ships must redirect their course to avoid hydropower systems.

9 Wind power

Wind (kinetic energy) is the movement of air across the surface of the Earth, affected by areas of higher pressure and of lower pressure. Wind power uses the wind as a primary source of energy. Wind power is the use of the wind to transfer energy and generate useful power. Therein, power generation (energy transfer) from wind ("wind energy") is based on the kinetic energy provided by air currents. It could also be considered mechanical energy, as kinetic and potential energy, because the energy is moving between two potentials (high and low pressure). Effectively, wind power is the harvesting of atmospheric air flow (air current) in order to transfer energy and produce power. In a wind turbine-electric system, for example, the [kinetic] energy of moving air (wind) is transferred into rotational mechanical energy (the wind rotates the blades of a turbine), which in turn produces electricity (the turbine turns a shaft rotating an electrical generator). In other words, the kinetic energy of the wind (air current) turns the blades of a turbine-type wind power generation system, which rotates a rotor that produces rotational mechanical power. In a turbine-electric system, the rotor connects to a shaft that leads to the electric power generator.

> **NOTE:** *Wind power, like hydropower, is also a manifestation of solar power.*

There are four primary types of wind power system (a.k.a., wind harvesting systems; wind energy conversion system, WECS) -- note that some of these are also propulsion systems:

1. Wind turbines (aerofoil-powered systems), which include:
   A. Windmills (e.g., gristmills that grind grain into flour and windpumps that move water) - produce mechanical and/or fluid power.
   B. Wind turbine-electric (a.k.a., wind-electric turbines, wind turbines, aerofoil powered generator) - produce electrical power.
2. Airborne wind power, which include:
   A. Kite propulsion - mechanical propulsion power
   B. Airborne wind power (non-propulsion) - kites/aircraft designed to produce mechanical and/or electrical power.
3. Sails (e.g., sailing boat sails) - mechanical propulsion power.
4. Magnus effect systems - mechanical propulsion power.

> **CLARIFICATIONS:** *The term, 'wind turbine', appears to have been adopted from hydroelectric technology (rotary propeller). The technical description of a wind turbine is an aerofoil-powered system/generator. A wind energy conversion system (WECS), or wind energy/power harvester is a machine that, powered by the movement of the wind, generates mechanical power that can be used to directly power machinery (mill, pump, etc.) or to power an electrical generator to produce electricity.*

The capacity of a wind powered device to harvest wind and transfer the energy to another carrier is affected by several parameters, including but not limited to:

1. Variability of wind at the site.
2. Aerodynamic design
3. Weight and positioning of blades (and other structural components)
4. Size of the generator relative to the turbine's swept area.

9.1 Wind power formula

The [wind] power generated by wind is:

- Pwi = $pAV^3/2$
  - where, p is air density (kg/m3), A is a projected area (m3), and V is wind speed (m/s).
  - Note: Wind power is abbreviated Pwi

Total wind energy flowing through an imaginary surface with area (A) during the time (t) is expressed as:

- $E = ½mv^2 = ½(Avt\rho)v^2 = ½At\rho v^3$
  - Where, ρ is the density of air; v is the wind speed; Avt is the volume of air passing through A (which is considered perpendicular to the direction of the wind); Avtρ is therefore the mass m passing through "A". Note that ½ ρv2 is the kinetic energy of the moving air per unit volume.
  - Power is energy per unit time, so the wind power incident on A (e.g. equal to the rotor area of a wind turbine) is:
- $P = E/t = ½At\rho v^3$
  - Wind power in an open air stream is thus proportional to the third power of the wind speed; the available power increases eightfold when the wind speed doubles. Wind turbines for grid electricity therefore need to be especially efficient at greater wind speeds.
  - The power in the wind of the area A, perpendicular to the wind direction, is given by the formula:
- $P = ½A\rho v^3$
  - The theoretical power available from wind:
  - The power P available in the wind impinging on a wind driven generator is given by:
- $P = ½CA\rho v^3$
  - Where, C is an efficiency factor known as the Power Coefficient which depends on the machine design, A is the area of the wind front intercepted

by the rotor blades (the swept area), ρ is the density of the air (averaging 1.225 Kg/m3 at sea level) and v is the wind velocity.

- Note that the power is proportional to area swept by the blades, the density of the air and to the cube of the wind speed. Thus doubling the blade length will produce four times the power and doubling the wind speed will produce eight times the power.
- Note also that the effective swept area of the blades is an annular ring, not a circle, because of the dead space around the hub of the blades.

NOTE: *According to Betz's law, the maximal achievable extraction of wind power by a wind turbine is 16/27 (59.3%) of the total kinetic energy of the air flowing through the turbine. Hence, the maximum theoretical power output of a wind machine is 0.59 times the kinetic energy of the air passing through the effective disk area of the machine.*

- If the effective area of the disk is A, and the wind velocity v, the maximum theoretical power output P is:
  - $P = E/t = 0.59½A\rho v^3$
    - Where, ρ is air density
    - The fraction of the energy captured by a wind turbine is given by a factor C_p, called the power coefficient and is defined as:
- $C_p(\lambda) = 16/27\eta_{turbine}(\lambda)$
  - Where, λ is the tip speed ratio of the blade, i.e. the tip speed divided by the wind speed and ηturbine is the efficiency of the turbine. Betz' law states that less than 16/27 (or 59%) of the kinetic energy in the wind can be converted to mechanical energy using a wind turbine. The power coefficient indicates how efficiently a turbine converts the energy in wind to electricity. Very simply, the electrical power output is divided by the wind energy input to measure how technically efficient a wind turbine is. The power curve divided by the area of the rotor gives the power output per square meter of rotor area.

9.1 Wind power system placement

Possible wind power locations include:

1. **Land/onshore** - refers to the construction of wind power systems on land. Conventional ground-level wind power is limited to the surface wind velocity.
2. **Offshore** - refers to the construction of wind turbines in a large body of water. In general, offshore wind is steadier and stronger than on land, and offshore wind arrays have less visual impact, but construction and maintenance requirements are higher.
3. **Infrastructural** - a wind turbine connected or somehow attached to infrastructure, such as the roof/side of a building or mast of a sailing boat.
4. **Airborne / atmospheric** - refers to the construction of wind turbines that float or are otherwise airbone like a kite or balloon. Airborne wind power can operate at a variety of heights above the earth's surface, and thus, has the potential to harvest higher-altitude winds, which become more consistent, predictable, and of a higher velocity at higher altitudes.

Surface and low-altitude wind power gives variable power which is very consistent from year to year, but which has significant variation over shorter time scales. Wind power can be highly variable at several different timescales: hourly, daily, or seasonally. Solar power tends to be complementary to wind. On a daily to weekly timescales, high pressure areas tend to bring clear skies and low surface winds, whereas low pressure conditions tend to be windier and cloudier. On seasonal timescales, solar energy peaks in summer, whereas in many areas wind energy is lower in summer and higher in winter. Thus, the intermittencies of wind and solar power tend to cancel each other somewhat. Similarly, conventional hydroelectricity complements wind power well. When the wind is blowing strongly, nearby hydroelectric stations can temporarily hold back their water. When the wind drops they can, provided they have the generation capacity, rapidly increase production to compensate.

NOTE: *The major drawback of surface and low altitude wind power is the variability of the wind. In some locations, some of the time, wind speed happens to be positively correlated with peak electricity use. In other locations on the planet, it is not positively correlated.*

When in an array, wind turbines must have sufficient space between them. On most horizontal wind turbine farms, a spacing of about 6-10 times the rotor diameter is often upheld. However, for large wind farms distances of about 15 rotor diameters should be more economically optimal, taking into account typical wind turbine and land costs. This conclusion has been reached by research conducted by Meneveau and Meyers (2012) based on computer simulations that take into account the detailed interactions among wind turbines (wakes) as well as with the entire turbulent atmospheric boundary layer. Moreover, recent research by John Dabiri of Caltech suggests that vertical wind turbines may be placed much more closely together so long as an alternating pattern of rotation is created allowing blades of neighbouring turbines to move in the same direction as they approach one another. (Calaf, 2010)

9.2 Wind supply characteristics

The wind has characteristics that should be accounted for in the design of wind powered devices. The wind always has a velocity and a direction, and wind powered devices must account for both (or they may not operate efficiently or safely). Many of the characteristics of wind are relative to the devices intended location of placement.

> **NOTE:** *Wind loads are cyclical because of natural variability in wind speed and wind shear (higher speeds at top of rotation).*

Site-specific wind characteristics relevant to wind powered devices include:

1. Wind specific characteristics:
    A. **Wind speed (wind flow velocity)** - caused by air moving from high pressure to low pressure, usually due to changes in temperature.
        - The Beaufort scale is one measure in common use for wind force and power (from 0 = calm to 12 = hurricane force).
    B. **Mean wind speed (average wind speed)** - the average speed of the wind over a period of time (e.g., annual). A data point, but does not tell how often "high" wind speeds occur.
        - Average wind speeds usually tend to increase with height then level off which is why wind turbines are usually installed as high above ground as possible.
        - Note that published average wind speeds are only reliable for open rural environments. Wind speeds just above roof level in urban environments will be considerably less than the quoted averages because of turbulence and shielding caused by buildings and trees.
    C. **Modal wind speed** - the speed at which the wind most frequently blows. A data point, but does not tell how often "high" wind speeds occur.
    D. **Wind speed distribution:** diurnal, seasonal, annual patterns.
    E. **Turbulence (turbulent flow)** - characterized by "chaotic" changes in pressure and flow. Measured as short-term fluctuations and long-term fluctuations. The wind itself may contain turbulence, and the affect of the rotating blades (and turbine) on the wind will generate their own turbulence.
    F. **Distribution of wind direction** - the frequency at which the wind blows with any particular speed follows a Rayleigh Distribution.
    G. **Wind shear (atmospheric windshear/wind gradient)** - variation in wind speed and/or direction over a relatively short atmospheric distance.
2. Atmospheric characteristics (atmospherics):
    A. **Air density (atmospheric density)** - the mass per unit volume of atmosphere.
    B. **Air pressure (atmospheric/barometric pressure)** - the pressure exerted by the weight of air in the atmosphere of Earth.
    C. **Temperature (atmospheric temperature)** - thermal quality of atmosphere.
    D. **Humidity (atmospheric humidity)** - the amount of water vapor in the air.
    E. **Air viscosity** - a measure of a fluid's resistance to gradual deformation by shear stress or tensile stress.

Wind velocities increase at higher altitudes due to surface aerodynamic drag (by land or water surfaces) and the viscosity of the air. The variation in velocity with altitude, called wind shear, is most dramatic near the Earth's surface. Typically, the variation follows the wind profile power law, which predicts that wind speed rises proportionally to the seventh root of altitude. Hence, doubling the altitude of a turbine, is expected to increase wind speeds by 10% and the expected power by 34%. Wind velocity increases logarithmically with the height above the earth's surface (altitude), reaching maximum velocity at 7 to 12 km above the surface.

The term power density is commonly used to compare geographical and altitudinal wind velocities and air densities. Power density is an expression of the potential power per turbine area that the wind possess in a certain location. This measure is useful because it does not depend on the turbine or power generator, but rather simply on the atmospheric conditions important to power production.

Although varying globally, the mean wind power density at 1km above the earth's surface is four times greater than a conventional wind turbine height (approximately 100 meters). And, at an altitude of 10 km, the power density is more than 40 times greater.

> **NOTE:** *In concern to surface-based wind power systems, in order to avoid buckling, doubling the tower height generally requires doubling the diameter of the tower as well, increasing the amount of material by a factor of at least four.*

At night time, or when the atmosphere becomes stable, wind speed close to the ground usually subsides whereas at turbine hub altitude it does not decrease that much or may even increase. As a result, the wind speed is higher and a turbine will produce more power than expected from the 1/7 power law: doubling the altitude may increase wind speed by 20% to 60%. A stable atmosphere is caused by radiative cooling of the surface and is common in a temperate climate: it usually occurs when there is a (partly) clear sky at night. When the (high altitude) wind is strong (a 10-meter wind speed higher than approximately 6 to 7 m/s) the stable atmosphere is disrupted because of friction turbulence and the atmosphere will turn neutral. A daytime atmosphere is either neutral (no net radiation; usually with strong winds and heavy clouding) or unstable (rising air because of ground heating—by the sun). Here again the 1/7 power law applies or is at least a good approximation of the wind profile.

9.2.1 Wind resource assessment

A **wind resource assessment** is the process by which site-specific wind power energy production is assessed. A wind resource assessment includes:

1. Wind resource maps - a map of estimated wind resources (a.k.a., wind atlas).
2. Wind and atmospheric measurements and trends.
3. Calculations:
    A. Correlations between on-site meteorological towers.
    B. Correlations between long-term weather stations and on-site meteorological towers.
    C. Vertical shear to extrapolate measured wind speeds to turbine hub height.
    D. Wind flow modeling to extrapolate wind speeds across a site.
    E. Estimated energy production using a wind turbine manufacturer's power curve.
    F. Application of energy loss factors applied to gross energy production:
    G. Wind turbine wake loss
    H. Wind turbine availability
    I. Electrical losses
    J. Blade degradation from ice/dirt/insects
    K. High/low temperature shutdown
    L. High wind speed shutdown
    M. Curtailments due to grid issues
    N. Atmospheric simulation modeling
    O. Wind flow modeling
    P. Wind farm modeling
    Q. Medium scale wind farm modeling
    R. Software applications
    S. Wind data management
    T. Wind data analysis

Energy "generation" will cut in (start) when wind reaches a lower limit m/s speed relative to the specific design of the wind powered device. Under that lower limit (relative to the specific design of the wind powered device), energy "generation" is not possible. This is known as a **lower operating limit**. Wind devices also have an upper limit. High wind speeds cause high rotation speeds and high stresses in the wind device, which can result in damage to the installation. To avoid these dangerous conditions, wind turbines are usually designed to cut out (stop) at upper limit wind speeds, either by braking or feathering the rotor blades allowing the wind to spill over the blades. This is known as an **upper operating limit**. Because of the limitations of the wind device and its generating system, and also upper speed limit at which the wind turbine can safely be used, it may capture less than the available wind energy.

For a given wind speed, the wind energy available also depends on the elevation of the wind powered device above sea level. As the density of air decreases with altitude, the wind energy density also decreases -- wind energy is proportional to air density. However, at the same time, the actual wind speeds tend to increase with increases in elevation above ground level. Since the wind energy is proportional to the cube of the wind speed (theoretical wind power), the net effect is that wind energy tends to increase with the height above ground level, even though wind energy density (as air density) decreases.

9.3 Wind power types

Wind power types include, but may not be limited to:

1. Windmill
2. Wind turbine
3. Sail power
4. Airborne power
5. Magnus power

9.3.1 Wind power type: Windmill

A windmill is a type of wind turbine whose final output is mechanical power, and doesn't produce electricity. Windmills convert wind power into rotational energy by means of vanes called sails or blades. The mechanical power generated from a windmill may be used for any number of mechanical processes, including but not limited to, historically, milling grain and pumping water. Thus, windmills are often called gristmills (grinds grain into flour) and windpumps (moves water).

Windmills consist of sails (blades), a tower structure that holds the sails, and internal machinery. Gears inside a windmill convey power from the rotary motion of the sails to a mechanical device. A wind turbine (wind-electric) is a windmill-like structure specifically developed to generate electricity. Windmills came first, but are categorically a type of wind turbine.

9.3.2 Wind power type: Wind turbine

Wind turbines are used to capture wind power -- wind drives the turbine. Technically a wind turbine is anything that captures wind energy the application of a turbine-like system. Both windmills and wind-electric turbines (a.k.a., turbine) are wind turbines because they both use a turbine to harvest wind power. However, in general, a wind-electric turbine is just called a "wind turbine", which can be confusing. The standard wind-electric turbine of today consists of a turbine, which has three blades that face into the wind such that the tubular steel or concrete tower is behind the turbine (downwind). A wind turbine (wind-electric) installation consists of the necessary systems needed to capture the wind's energy, point the turbine into the wind, convert mechanical rotation into electrical power, and other systems to start, stop, and control the turbine.

An array of wind turbines (turbine + generator) is known as a wind farm. In other words, a wind farm is a group of wind turbines in the same location connected to an electric power transmission network.

Wind turbine power depends on both rotor speed and wind speed (as well as the turbine-electric system itself), and harvested power can be represented on a three-dimensional surface (with output power kW as the y-axis and wind speed m/s as the x-axis). The power that wind produces through a turbine is dependent on the area the turbine's blades cover as they sweep through the air, and also, the wind's velocity as it flows over the blades. To increase the power produced by a wind turbine, either the length of the turbine blades or the wind velocity must increase. Wind power generation is cubically proportional to wind velocity, while only linearly proportional to area. Therefore, doubling the size of a turbine's sweeping area would only double the power generated, but doubling the wind velocity flowing into the turbine will increase the power eight times.

9.3.2.1 Wind Turbine structural classifications

Wind turbines can rotate about a horizontal or a vertical axis, the former being both older and more common. They can also include blades (transparent or not) or be bladeless.

1. **Horizontal-axis wind turbines (HAWT):** The main rotor shaft and electrical generator at the top of a tower, and must be pointed into the wind. Small turbines are pointed by a simple wind vane, while large turbines generally use a wind sensor coupled with a servo motor. Most have a gearbox (for stepping up the speed), which turns the slow rotation of the blades into a quicker rotation that is more suitable to drive an electrical generator. Turbine towers produce turbulence (mast wake) behind them; hence, the turbine is usually positioned upwind of its supporting tower. Instead of a gearbox some turbine designs use the direct drive of an annular generator. Conventional horizontal axis turbines can be divided into three components:
   A. The rotor component: includes the blades for converting wind energy to low speed rotational energy.
   B. The generator component: includes the electrical generator, the control electronics, and most likely, a gearbox (e.g., planetary gearbox) and adjustable-speed drive or continuously variable transmission for converting the low speed incoming rotation to high speed rotation.
   C. The structural support component: includes the tower and rotor yaw mechanism.
   D. The turbine in a HAWT, also called "low-speed rotor", usually has two to six blades. The most common number of blades is three since they can be positioned symmetrically (120° apart), and maintain the system's lightness in mass, while ensuring the stability of the overall wind power system (WPS).
   E. Wind is a form of "linear" kinetic energy, and hence, horizontal axis turbines with a horizontally positioned shaft, ease the conversion of the wind's linear energy into a rotational one.
2. **Vertical-axis wind turbines (or VAWTs):** The main rotor shaft is arranged vertically. One advantage of this arrangement is that the turbine does not need to be pointed into the wind to be effective, which is an advantage on a site where the wind direction is highly variable. It is also an advantage when the turbine is integrated into a building because it is inherently less steerable. Also, the generator and gearbox can be placed near the ground, using a direct drive from the rotor assembly to the ground-based gearbox, improving accessibility for maintenance. The key disadvantages include: the relatively low rotational speed with a consequential higher torque; the inherently lower power coefficient; the 360-degree rotation of the aerofoil within the wind flow during each cycle leads to higher dynamic loading on the blade(s); the pulsating torque generated by some rotor designs on the drive train; and the difficulty of modelling the wind flow accurately and hence the challenges of analysing and designing the rotor prior to fabricating a prototype. In general, this design involves blades extending upwards that are supported by a rotating framework. Subtypes of the vertical axis design include:
   A. Darreius wind turbine
   B. Giromill
   C. Savonius wind turbine
   D. Twisted savonius
   E. Vortexis
3. Unconventional designs: These designs differ significantly from the most common types in use.
   A. Modified horizontal - for which there are many subtypes. Sub-types include, but are not limited to: twin-bladed rotor; downwind rotor; ducted rotor; co-axial, multi-rotor; counter-rotating horizontal-axis; furling tail and twisting blades; wind-mill style; ducted 2-blade HAWT.
   B. Modified vertical axis - for which there are many subtypes. Sub-types include, but are not limited

to: aerogenerator; savonius; augmented.

C. VAWTs are not self-starting machines and must be started in motoring mode, and then switched to generating mode.
D. Aerial - airborne wind turbines; high-altitude wind power; crosswind kite power.
E. Blade Tip Power System (BTPS)
F. Fuller - The “Fuller” wind turbine is a fully enclosed wind turbine that uses boundary layers instead of blades.
G. H-rotor - one blade is pushed by the wind while the other is being pushed in the opposite direction. Consequently, only one blade is working at a time.
H. INVELOX - not a turbine, rather a wind capturing and delivery system to a turbine.
I. Motion-driven - drive by the motion of objects (e.g., cars) moving past.
J. Piezoelectric - Turbines with diameters on the scale of 10 centimeters work by flexing piezoelectric crystals as they rotate.
K. Ram air turbine (RAT) - a turbine fitted to small aircraft.
L. Saphonian - uses a dish to generate wind pressure and back-and-forth motion that drives a piston.
M. Solar chimney - Wind turbines may also be used in conjunction with a solar collector to extract the energy due to air heated by the Sun and rising through a large vertical Solar updraft tower.
N. Vaneless ion wind generator - produces electrical energy directly by using the wind to pump electric charge from one electrode to another, with no moving parts.
O. Vortex bladeless - The vortex bladeless device deliberately maximizes vortex shedding, converting wind energy to fluttering of a lightweight vertical pole, then captures that energy with a generator at the bottom of the pole.
P. Windbeam - The generator consists of a lightweight beam suspended by durable long-lasting springs within an outer frame. The beam oscillates rapidly when exposed to airflow due to the effects of multiple fluid flow phenomena. A linear alternator assembly converts the oscillating beam motion into usable electrical energy. A lack of bearings and gears eliminates frictional inefficiencies and noise.
Q. Wind belt - A tensioned but flexible belt vibrates by the passing flow of air, due to aeroelastic flutter. A magnet, mounted at one end of the belt translates in and out of coiled windings producing electricity.
R. Wind tower technology - A Wind Tower uses pressure differentials produced by wind flow around a building moving through a ducted turbine to generate electricity. A windcatcher assembly directs the flow into the tower, The tower structure together with the embedded nozzles inside it will accelerate the flow.

Advantage and disadvantage comparison:

1. VAWTs’ electrical machines and gearbox can be installed at the bottom of the tower, on the ground, whereas in HAWTs, these components have to be installed at the top of the tower, which requires additional stabilizing structure for the system.
2. Another advantage of the VAWTs is that they do not need the yaw mechanism since the generator does not depend on the wind direction.

9.3.2.2 Wind turbine electric generators

Turbine generators turn the shaft power of the turbine into electric power. The generator in a wind turbine produces alternating current (AC) electricity, which may be rectified to produce DC. However, wind turbines can connect to both AC and DC grids. When connected to an AC grid, some turbines have an AC>DC>AC converter—which converts the AC to DC, with a rectifier, and then back to AC with an inverter, in order to match the frequency and phase of the grid. However, the most common method in large modern turbines is to use a doubly fed induction generator directly connected to the electricity grid.

Modern turbines use variable speed generators combined with a partial- or full-scale power converter between the turbine generator and the collector system, which generally have more desirable properties for grid interconnection and have low voltage ride through-capabilities.

As an AC generator speeds up and slows down, due to changes in the source of supplied energy (e.g., changes in the wind), the electrical output characteristics of the generator (its frequency and voltage) will change. Hence, either the rotor’s power output must be controlled, or the voltage and frequency output of the generator itself must be transformed/controlled.

Different types of wind turbine generators behave differently during transmission grid disturbances, so extensive modelling of the dynamic electromechanical characteristics of a new wind farm is required to ensure predictable stable behaviour during system faults.

In concern to power generation per rotation speed (i.e., speed architecture), wind turbines can operate at either fixed or variable speeds:

1. **Fixed-speed turbine/generator** (FS architecture; FSG) - turbine speed is not adjustable as a function

of wind speed; a fixed-rpm machine. These turbines have synchronous electric generators and operate on the grid's frequency. These machines are not the best solution for the wind turbines, because the wind always changes its speed. For fixed-rpm wind turbines, there is only one wind velocity on the turbine's power curve (power versus wind speed) at which the tip-speed ratio is optimum. Those turbines designed to operate at fixed speed (an FS architecture) generally have a gearbox in between the turbine's rotor and the shaft of the electric generator, and are connected directly to the grid as long as the wind speed is within operation limits. Constant speed using a synchronous generator (out-dated design). In other words, the generator expects a constant shaft speed. Fixed speed induction generators (FSIG) operate within a few percent of constant speed. When connected to an AC grid, fixed speed (FS) turbines use synchronous machines, and operate at an FS that depends on the grid's frequency. Early versions of the wind turbine were fixed speed turbines; that is, the rotor speed was a constant for all wind speeds.

2. **Variable-speed turbine/generator** (VS architecture; adjustable speed generator; ASG) - turbine speed is adjusted as a function of wind speed. These turbines can operate over a wide-range of speeds and usually have asynchronous generators. When the input power changes/fluctuates (randomly), then the output voltage and frequency is variable. In variable-speed turbines have greater output efficiency over fixed-speed turbines. A wind turbine that can produce power over a continuous range of rotor speeds (varying wind supply) can be controlled to operate constantly at or near its optimum tip-speed ratio. All variable-speed turbines require power electronics to change varying AC power to a constant voltage and frequency. Electrical distribution grids (to which a wind turbine may be connected) must maintain steady frequency and voltage levels to avoid damaging demand-side equipment of other users on the same network. Electrical harmonics are also a critical issue for any variable-speed design. Harmonics distort the normally smooth sinusoidal variation of a grid's voltage.
3. Since the input voltage and input frequency are variable, this variable AC magnitude and frequency must be converted to constant frequency AC.
   A. A variable-speed system may base its variability on generator speed variability and/or mechanical speed variability. Variable-speed generator methods/architectures are based on allowing the speed of the generator to vary as the supplied [wind] energy varies. Variable-speed mechanical methods are based on the use of continuously variable-speed mechanical or hydraulic drives, which allow the rotor rpm to vary while maintaining a constant generator speed. By using power electronics and controllers, AC can be converted to DC and then back to AC (AC>DC>AC) to produce a reliable and steady frequency, instead of a just outputting varying AC voltages and frequencies ("wild AC"). In other words, the "wild" AC is rectified into a steady direct current, which is then inverted to grid-grade alternating current of constant voltage and frequency. Variable-speed operation was accomplished with a current source-load commutated inverter, also known as a DC current link frequency converter. This provided AC-DC-AC conversion These power control elements can be located within the turbine itself, or located at some distance away from the turbine where its output connects to an electrical circuit/grid.
   B. Synchronous generator with in-line frequency control - the rotor and turbine can be run at a variable speed corresponding to the prevailing wind conditions. This will produce a varying frequency output from the generator synchronised with the drive shaft rotation speed. This output can then be rectified in the generator side of an AC-DC-AC converter and the converted back to AC in an inverter in grid side of the converter which is synchronised with the grid frequency.
   C. Doubly fed induction generator (DFIG) - the DFIG system consists of a 3 phase wound rotor generator with its stator windings fed from the grid and its rotor windings fed via a back to back converter system in a bidirectional feedback loop taking power either from the grid to the generator or from the generator to the grid. The doubly fed induction generator design means that the electronic control circuits and frequency converter do not have to be dimensioned to carry the full generator power.
   D. Variable speed turbines use DC machines, brushless DC (BLDC) machines, and induction machines. DC machines are not commonly used due to the maintenance problems with the brushes.
   E. AC regulation by the control system: The control system regulates the speed of the blades, and hence torque of the shaft, in an effort to match

the electrical networks required parameters. This may be done by changing the pitch of the blade tips.

F. Generator operating principle: The feedback control system monitors the stator output voltage and frequency and provides error signals if these are different from the grid standards. The frequency error is equal to the generator slip frequency and is equivalent to the difference between the synchronous speed and the actual shaft speed of the machine.
G. Grid Side Converter (GSC): Carries current at the grid frequency.
H. Machine Side Converter (MSC): Carries current at slip frequency. It is a DC to AC inverter used to provide variable AC voltage and frequency to the rotor to control torque and speed.

> **NOTE:** *The output of a variable-speed turbine with no frequency/voltage output controls is known as "***wild AC***", because it varies with the variability of the "wild" primary energy source.*

A typical fixed speed system employs a rotor with three variable pitch blades (generally), which are controlled automatically to maintain a fixed rotation speed for any wind speed. The rotor drives a synchronous generator through a gear box, and the whole assembly is housed in a nacelle on top of a substantial tower with massive foundations requiring hundreds of cubic metres of reinforced concrete. Fixed speed systems may suffer excessive mechanical stresses, because they are required to maintain a fixed speed regardless of the wind speed. There is no "give" in the mechanism to absorb gusty wind forces, resulting in high torque, high stresses and excessive wear on the gear box, increasing maintenance requirements and reducing service life.

Variable speed wind turbines can capture more of the wind's energy than constant speed machines; they can speed up and slow down per wind conditions, and the electronic control systems will keep the generator's output frequency constant during fluctuating wind conditions. For variable speed wind turbines, one of two types of generators can be used: a DFIG (doubly fed induction generator) or an FRC (fully rated converter). For variable speed wind turbines, one of two types of generators can be used: a DFIG (doubly fed induction generator) or an FRC (fully rated converter).

A DFIG generator draws reactive power from the electrical network; this can increase the vulnerability of a transmission system in the event of a failure. A DFIG configuration will require the generator to be a wound rotor; squirrel cage rotors cannot be used for such a configuration.

Consider a variable speed wind turbine with a permanent magnet synchronous generator. The generator produces AC electricity. The frequency of the AC voltage generated by the wind turbine is a function of the speed of the rotor within the generator:

- N = 120f/P
- where N is the rotor speed, P is the number of poles in the generator, and f is the frequency of the output Voltage. That is, as the wind speed varies, the rotor speed varies, and so the frequency of the Voltage varies. This form of electricity cannot be directly connected to an AC balanced transmission system. Instead, its AC electrical output must be corrected such that its frequency is constant. For this, power converters are employed, which results in the de-coupling of the wind turbine from the transmission system. As more wind turbines are included in a national power system, the "inertia" of the transmission system is decreased. This means that the frequency of the transmission system is more strongly affected by the loss of a single generating unit.

The voltage generated by a variable speed wind turbine is non-grid compliant (AC grid). In order to supply the transmission network with power from these turbines, the signal must be passed through a power converter, which ensures that the frequency of the voltage of the electricity being generated by the wind turbine is the frequency of the transmission system when it is transferred onto the transmission system. Power converters first convert the signal to DC, and then convert the DC signal to an AC signal.

> **NOTE:** *All wind turbines that generated electricity were variable speed before 1939. All grid-connected wind turbines, from the first one in 1939 until the development of variable-speed grid-connected wind turbines in the 1970s, were fixed-speed wind turbines. As of 2003, nearly all grid-connected wind turbines operate at exactly constant speed (synchronous generators) or within a few percent of constant speed (induction generators). (Bassyouni, 2013)*

9.3.2.3 Wind turbine aerodynamics

Wind powered devices are designed [in part] for their aerodynamic characteristics. Aerodynamics the study of the properties of moving air, and the movement of substances and system moving through air -- the interaction between the air and solid bodies moving through it. As the wind blows, an aerodynamic force produces a torque that is transmitted through the drive-train to the generator.

The aerodynamic efficiency of a wind turbine depends on the wind's characteristics and the design of the wind turbine itself:

1. Design of the airfoils (blades) - The aerodynamics of the blades/airfoils (wing-shaped) include: chord

length; blade shape; blade mount position/angle - angle of attack; blade dimensions; and revolutions per minute (rpm). Note that aerodynamic torque [from wind] is "captured" by the blades. The design of the airfoil and its angle of attack are critical to the power-producing capacity of the rotor. Each airfoil/blade has an optimum angle of attack to produce an optimum lift-to-drag ratio (the point at which the airfoil will have its optimum performance).
2. Design of the system's other structural components: rotor hub (if there is one), nacelle, tower structure, and foundation.

In addition to aerodynamic design of the blades, the design of a complete wind power system must also address the design of the hub, controls, generator, supporting structure and foundation. Further design questions arise when integrating wind turbines into electrical power grids.

> **NOTE:** *It is generally understood that noise increases with higher blade tip speeds.*

9.3.2.4 Wind turbine blades/airfoils

The number of blades in the turbine's rotor, their aerodynamic design, and the rotor's rotational speed must be optimised to extract the maximum amount of power (energy) from the available wind. While using rotors with multiple blades may capture more wind energy, there is a practical limit to the number of blades that can be used, because each blade of a spinning rotor leaves turbulence in its wake and reduces the amount of power the following [rotational] blade can extract from the wind. This same turbulence effect also limits the possible rotor speed, because a high speed rotor does not provide enough time for the air flow to settle after the passage of a blade before the next blade comes along.

There is a lower limit to both the number of blades and the rotor speed for the turbine to function. With too few rotor blades, or a slow turning rotor, most of the wind will pass undisturbed through the gap between the blades reducing the potential for capturing the wind energy. The fewer the number of blades, the faster the wind turbine rotor needs to turn to extract maximum power from the wind.

The ratio between the speed of the blade tips and the speed of the wind is called **tip-speed ratio (TSR)**. In part, tip-speed ratio is a concept used by wind turbine designers to optimise a blade set to the shaft speed required by a particular electricity generator to generate maximum energy output from the available wind. It is also one way of comparing performance between variable- and constant-speed turbine operation.

Tip-speed-ratio is generally plotted on a two dimensional graph with wind speed (m/s) as the x axis, and rotor speed (rpm) as the y axis.

Operating closer to their optimal tip speed ratio during energetic gusts of wind allows wind turbines to improve energy capture from sudden gusts (typical in urban settings).

The tip-speed-ratio for a wind turbine is given by the following formula (λ, lambda = TSR):

- $\lambda = \omega R/v$ (sometimes written as $TSR=\Omega R/V$)
  - Where, ω is the rotor speed (i.e., the angular velocity of the rotor; in radians per second), R is the length of a blade (the distance between the axis of rotation and the tip of the blade), and v is the wind speed.
- For a fixed-speed wind turbine, the value of the tip-speed ratio is only changed by wind speed variations. In reference to a C_p-λ graph (the power coefficient versus tip-speed ratio C_p vs λ curve), illustrates the relationship between tip-speed ratio and efficiency, it is evident that only one value of λ yields the highest efficiency. C_p is the power co-efficient. That is, the fixed speed wind turbine is not operating at peak efficiency across a range of wind speeds.

The tangential velocity S of any blade section at a distance r from the centre of rotation (the root of the blade) is given by:

1. $S = r\,\Omega$
2. Where, Ω is the angular velocity of rotation in radians.

Larger rotor blades are useful for maximizing air stream conversion to mechanical energy. Unfortunately, larger blades (in high wind conditions) lead to very high tip speeds. And, higher tip speeds equate to higher (possibly unacceptable) noise levels. Hence, depending upon the surrounding environment smaller blades/turbines may be necessary to reduce noise levels to a sufficiently safe (and non-polluting/non-disturbing) level.

Additionally, system reliability is affected by blade count and weight through the dynamic loading of the rotor within the drive train and tower systems. While aligning the wind turbine to changes in wind direction (yawing), each blade experiences a cyclic load at its root end depending on blade position. This is true of one, two, three blades or more. When these loads are symmetrical, the turbine with yaw more smoothly during operation. Turbines with one or two blades can use a pivoting teetered hub to also nearly eliminate the cyclic loads into the drive shaft and system during yawing.

> **NOTE:** *High capacity wind turbines, such as those used by the electricity utilities in the electricity grid, typically have blades with a cross section similar to the aerofoils used to provide the lift in aircraft wings.*

In general, ideal materials for blades should meet the following criteria:

1. Low weight or density to reduce gravitational forces.
2. High [tensile] strength to withstand strong loading of wind and gravitational force of the blade itself.
3. High fatigue resistance to withstand cyclic loading.
4. High stiffness to ensure stability of the optimal shape and orientation of the blade and clearance with the tower.
5. High fracture toughness.
6. The ability to withstand environmental impacts such as lightning strikes, humidity, and temperature.

Today, wind turbine blades are mainly made of composite, materials including: polyester resin, vinyl resin, epoxy thermosetting matrix resin, E-glass fibers, S-glass fibers, and carbon fiber reinforced materials. Construction may use manual layup techniques or composite resin injection molding. The majority of current commercialized wind turbine blades are made from fiber-reinforced polymers (FRP's), which are composites consisting of a polymer matrix and fibers. The long fibers provide longitudinal stiffness and strength, and the matrix provides fracture toughness, delamination strength, out-of-plane strength, and stiffness. Material indices based on maximizing power efficiency, and having high fracture toughness, fatigue resistance, and thermal stability, have been shown to be highest for glass and carbon fiber reinforced plastics (GFRP's and CFRPs). (Griffin, 2003)

Use of aluminum and composite materials in their blades has contributed to low rotational inertia, which means that newer wind turbines can accelerate quickly if the winds pick up, keeping the tip speed ratio more nearly constant.

Manufacturing blades in the 40 to 50 metre range involves fibreglass composite fabrication techniques. There are a variety of applied variations on this technique, some including carbon and wood with fibreglass in an epoxy matrix. Other options include pre-impregnated ("prepreg") fibreglass and vacuum-assisted resin transfer molding. Each of these options use a glass-fibre reinforced polymer composite constructed with differing complexity. Perhaps the largest issue with more simplistic, open-mould, wet systems are the emissions associated with the volatile organics released. Pre-impregnated materials and resin infusion techniques avoid the release of volatiles by containing all VOC's. However, these contained processes have their own challenges, namely the production of thick laminates necessary for structural components becomes more difficult. As the preform resin permeability dictates the maximum laminate thickness, bleeding is required to eliminate voids and ensure proper resin distribution.

NOTE: *The transportation of long blades and towers to their final placement may require unconventional transportation methods due to their significant length.*

9.3.2.5 Wind turbine control systems

The control system of a wind turbine allows for programmed adjustment of the system (mechanically and/or electrically) for maximum, or less than maximum, power output. Wind turbines designed for maximum power output (i.e., maximum power point tracking) attempt to pull the maximum possible electrical power from a given turbine under the current wind conditions. However, wind turbines control systems can also be designed to deliberately pull less electrical power than they possibly could in most circumstances, in order to provide other benefits, which include:

1. Spinning reserves to quickly produce more power when needed—such as when some other generator suddenly drops from the grid—up to the max power supported by the current wind conditions.
2. Variable-speed wind turbines can (very briefly) produce more power than the current wind conditions can support, by storing some wind energy as kinetic energy (accelerating during brief gusts of faster wind) and later converting that kinetic energy to electric energy (decelerating, either when more power is needed elsewhere, or during short lulls in the wind, or both).
3. Damping (electrical) subsynchronous resonances in the grid.
4. Damping (mechanical) resonances in the tower.

In wind turbines with a nacelle, the nacelle houses the gearbox and generator connecting the tower and rotor. Sensors detect the wind speed and direction, and motors turn the nacelle into the wind to maximize output. In conventional wind turbines, the blades spin a shaft that is connected through a gearbox to the generator. When present, a gearbox is inserted between the rotor hub and the generator. The gearbox increases the generator's incoming shaft rotations per minute. The presence of a gearbox may allow a reduction in the generators weight. In gearless wind turbines (also called 'direct drive turbines') the rotor shaft is attached directly to the generator, which spins at the same speed as the blades.

Wind turbines without gearboxes are called direct-drive wind turbines. An advantage of a gearbox is that generators are typically designed to have the rotor rotating at a high speed within the stator. Direct drive wind turbines do not exhibit this feature.

NOTE: *Older style wind turbines rotated at a constant speed, to match power line frequency, which allowed the use of less costly induction generators[citation needed]. Newer wind turbines often turn at whatever speed generates* *electricity most efficiently. The varying output frequency and voltage can be matched to the fixed values of a power network (the "grid") using multiple technologies (e.g., doubly fed induction generators or full-effect converters) where the variable frequency current produced is converted to DC and then back to AC.*

The speed at which a wind turbine rotates must be controlled for efficient power generation and to keep the turbine components within designed speed and torque limits. The centrifugal force on the spinning blades increases as the square of the rotation speed, which makes this structure sensitive to overspeed. If the rated wind speed is exceeded the power has to be limited via a control system. A control system involves three basic elements: sensors to measure process variables, actuators to manipulate energy capture and component loading, and control algorithms to coordinate the actuators based on information gathered by the sensors.

NOTE: *Wind powered devices often include an anemometer for measuring wind speed and direction (a common weather station instrument).*

Methods of control for dumping power include (because of a rapid increase in the velocity of the wind):

1. Change aerodynamic efficiency
   A. Variable pitch, feather or stall
   B. Operate at constant rpm
   C. Spoilers
2. Change intercept area
   A. Yaw rotor out of wind
   B. Change rotor geometry
3. Brake
   A. Mechanical, hydraulic
   B. Air brake
   C. Electrical (resistance, magnetic)

9.3.2.6 Wind turbine pitch control

Active pitch control, where the blades are twisted from the hub. The pitch angle of turbine blades can be changed to control for speed (blade control):

1. Stalling works by increasing the angle at which the relative wind strikes the blades (angle of attack), and it reduces the induced drag (drag associated with lift). Stalling is simple because it can be made to happen passively (it increases automatically when the winds speed up), but it increases the cross-section of the blade face-on to the wind, and thus the ordinary drag. A fully stalled turbine blade, when stopped, has the flat side of the blade facing directly into the wind.
2. Furling works by decreasing the angle of attack, which reduces the induced drag from the lift of the rotor, as well as the cross-section. One major problem in designing wind turbines is getting the blades to stall or furl quickly enough should a gust of wind cause sudden acceleration. A fully furled turbine blade, when stopped, has the edge of the blade facing into the wind. Since furling requires acting against the torque on the blade, it requires some form of pitch angle control, which is achieved with a slewing drive.
3. Some blades are designed to automatically increase their angle of attack at higher wind speed, as the blades speed up.

Modern large wind turbines are variable-speed machines. When the wind speed is below rated, generator torque is used to control the rotor speed in order to capture as much power as possible. Applied generator torque controls the rotor speed.

NOTE: *All turbines are equipped with protective features to avoid damage at high wind speeds, by feathering the blades into the wind which ceases their rotation, supplemented by brakes.*

There are three ways of braking (slowing) the rotation of a turbine: aerodynamic stalling or furling, electrical breaking (electromagnetic breaking), and/or mechanical breaking. In concern to electrical breaking, the braking of a small wind turbine can be done by dumping energy from the generator into a resistor bank, converting the kinetic energy of the turbine rotation into heat. This method is useful if the kinetic load on the generator is suddenly reduced or is too small to keep the turbine speed within its allowed limit. Doing this process cyclically (i.e., cyclically braking) causes the blades to slow down, which increases the stalling effect, reducing the efficiency of the blades. This way, the turbine's rotation can be kept at a safe speed in faster winds while maintaining (nominal) power output. In concern to mechanical breaking, a mechanical drum brake or disk brake is used to stop turbine in emergency situation such as extreme gust events or over speed. This brake is a secondary means to hold the turbine at rest for maintenance, with a rotor lock system as primary means. Such brakes are usually applied only after blade furling and electromagnetic braking have reduced the turbine speed generally 1 or 2 rotor RPM, as the mechanical brakes can create a fire inside the nacelle if used to stop the turbine from full speed. The load on the turbine increases if the brake is applied at rated RPM. Mechanical brakes are driven by hydraulic systems and are connected to main control box.

9.3.2.7 Wind turbine yaw control

Wind devices that operate with their blades or sales pointing into the wind require a control systems for reorienting their blades/sales into the wind as the wind direction changes (typically, horizontal axis wind turbines). In general, only the yaw axis is used for this type of reorientation (although, the blades themselves may also have secondary re-orientational capabilities). Hence, the 'yaw system' of wind turbines is the component responsible for the orientation of the wind turbine rotor into the wind.

By minimizing the yaw angle (the misalignment between wind and turbine pointing direction), the power output is maximized and non-symmetrical loads minimized. However, since the wind direction varies quickly the turbine will not strictly follow the direction and will have a small yaw angle on average. The power output losses can simply be approximated to fall with $(\cos(\text{yaw angle}))^3$. Particularly at low-to-medium wind speeds, yawing can make a significant reduction in turbine output, with wind direction variations of ±30° being quite common and long response times of the turbines to changes in wind direction. At high wind speeds, the wind direction is less variable.

There are two types of yaw system:

1. Passive yaw (self-orientation)
2. Active yaw (automatic mechanical orientation)

Both passive and active systems require a yaw bearing, which allows for a horizontal rotation of the turbine element itself. In active systems, in order to stabilize the yaw bearing against rotation a means of braking is necessary. Active systems involve yaw drives (consisting of an electric motor and gearbox), and passive systems involve a yaw vane.

Passive yaw systems utilize the wind's force itself in order to adjust the orientation of the wind turbine rotor into the wind. In their simplest form these system comprise a roller bearing connection between the tower and the nacelle and a tail fin (yaw vane) mounted on the nacelle, designed in such a way that it turns the wind turbine rotor into the wind by exerting a "corrective" torque to the nacelle. Therefore, the power of the wind is responsible for the rotor rotation and the nacelle orientation. The tail fin (or yaw vane) is commonly used for small wind turbines since it offers a low cost and reliable solution. It is however unable to cope with the high moments required to yaw the nacelle of a large wind turbine.

Alternatively in case of downwind turbines, the tail fin is not necessary since the rotor itself is able to yaw the nacelle into the wind.

Passive yaw systems have to be designed in a way that the nacelle does not follow the sudden changes in wind direction with too fast a yaw movement, in order to avoid high gyroscopic loads. Additionally the passive yaw systems with low yaw-friction are subjected to strong dynamic loads due to the periodic low amplitude yawing caused by the variation of the inertia moment during the rotor rotation. This effect becomes more severe with the reduction of the number of blades.

Active yaw systems use a torque producing device for rotating the nacelle of the wind turbine against the stationary tower, based on automatic signals from wind direction sensors or manual actuation (control system override). The design of an active yaw system varies depending on the design characteristics of the wind device; however, all active yaw systems include: a rotatable connection between nacelle and tower (yaw bearing); an active variation of the rotor orientation (i.e. yaw drive); a means of restricting the rotation of the nacelle (yaw brake) and a control system which processes the signals from wind direction sensors (e.g. wind vanes) and gives the proper commands to the actuating mechanisms.

Modern large wind turbines are typically actively controlled to face the wind direction measured by a wind vane situated on the back of the nacelle. This process is known as yawing.

9.3.2.8 Wind turbine design limits

For safety and efficiency reasons wind turbines are subject to operating limits depending on the wind conditions and the system design.

1. **Cut-in Wind Speed:** This is the minimum wind velocity below which no useful power output can be produced from wind turbine, typically between 3 and 4 m/s (10 and 14 km/h, 7 and 9 mph).
2. **Rated Wind Speed** (also associated with the Nameplate Capacity): This is the lowest wind velocity at which the turbine develops its full power. This corresponds to the maximum, safe electrical generating capacity which the associated electrical generator can handle, in other words the generator's rated electrical power output. The rated wind speed is typically about 15 m/s (54 km/h, 34 mph) which is about double the expected average speed of the wind. To keep the turbine operating with wind speeds above the rated wind speed, control systems may be used to vary the pitch of the turbine blades, reducing the rotation speed of the rotor and thus limiting the mechanical power applied to the generator so that the electrical output remains constant. Though the turbine works with winds speeds right up to the cut-out wind speed, its efficiency is automatically reduced at speeds above the rated speed so that it captures less of the available wind energy in order to protect the generator. While it would be possible to use larger generators to extract full power from the wind at speeds over the rated wind speed, this would not normally be economical because of the lower frequency of occurrence of wind speeds above the rated wind speed.
3. **Cut-out Wind Speed:** This is the maximum safe working wind speed and the speed at which the wind turbine is designed to be shut down by applying brakes to prevent damage to the system.

In addition to electrical or mechanical brakes, the turbine may be slowed down by stalling or furling.

A. Stalling: This is a self correcting or passive strategy which can be used with fixed speed wind turbines. As the wind speed increases so does the wind angle of attack until it reaches its stalling angle at which point the "lift" force turning the blade is destroyed. However increasing the angle of attack also increases the effective cross section of the blade face-on to the wind, and thus the direct wind force and the associated stress on the blades. A fully stalled turbine blade, when stopped, has the flat side of the blade facing directly into the wind.

B. **Furling or Feathering:** This is a technique derived from sailing in which the pitch control of the blades is used to decrease the angle of attack which in turn reduces the "lift" on the blades as well as the effective cross section of the aerofoil facing into the wind. A fully furled turbine blade, when stopped, has the edge of the blade facing into the wind reducing the wind force and stresses on the blade.

The cut-out speed is specified to be as high possible consistent with safety requirements and practicality in order to capture as much as possible of the available wind energy over the full spectrum of expected wind speeds (See diagram of Wind Speed Distribution below). A cut-out speed of 25 m/s (90 km/h, 56 mph) is typical for very large turbines.

4. **Survival Wind Speed:** This is the maximum wind speed that a given wind turbine is designed to withstand above which it can not survive. The survival speed of commercial wind turbines is in the range of 50 m/s (180 km/h, 112 mph) to 72 m/s (259 km/h, 161 mph). The most common survival speed is 60 m/s (216 km/h, 134 mph). The safe survival speed depends on local wind conditions is usually regulated by national safety standards.

9.3.2.9 Wind turbine monitoring

Wind turbines must be monitored for their structural health/safety and performance, which requires data transmission. Structural monitoring is usually done through accelerometers and strain gages attached to the nacelle to monitor the gearbox and other equipment. Digital image correlation and stereophotogrammetry are used to measure dynamics of wind turbine blades. These methods usually measure displacement and strain to identify locations of defects. Dynamic characteristics of non-rotating wind turbines have been measured using digital image correlation and photogrammetry. Three dimensional point tracking has also been used to measure rotating dynamics of wind turbines

Wind turbines require regular maintenance to stay operable (operability - reliable and available). Modern turbines usually have a small onboard crane for hoisting maintenance tools and minor components. However, large heavy components like generator, gearbox, blades and so on are rarely replaced and a heavy lift external crane is needed in those cases. If the turbine has a difficult access road, a containerized crane can be lifted up by the internal crane to provide heavier lifting. Wind turbines can be "repowered", meaning that instead of installing new turbines, an existing turbine is replaced with a larger and more powerful one.

A major cause of wind turbine failure is the accumulation of fatigue damage from turbine rotor fatigue loads (i.e., rapid changes in rotor torque due to wind gusts producing load spikes).

9.3.3 Wind power type: Sail power

A sail is means for redirecting the power of the wind to propel a craft on water, ice, or land. Hence, sails provide mechanical propulsion power to a traveling craft/ platform. In doing so, sails behave aerodynamically like wings, generating high and low pressure on either side of the sail, and hence, craft. Therein, the sails mobilize lift aerodynamics (as air passes along the surfaces), and they mobilize drag aerodynamics to the degree that air is directed at the surface. When both lift and drag are present, a sail will function similarly to a wing in a vertical orientation. Lift aerodynamics refers to the air moving faster and having a longer way to travel along the outside curve of the sail, and a slower, shorter travel along the inner curve.

In most cases sails are supported by a mast rigidly attached to the sailing craft, however some craft employ a flexible mount for a mast. Sails also employ spars and battens to determine shape in the axis perpendicular to the mast. As a result, sails come in a variety of shapes that include both triangular and quadrilateral configurations, usually with curved edges that promote curvature of the sail.

Sails propel the craft in one of two ways. When the craft is going in the direction of the wind (i.e., downwind), the sails may be set merely to trap the air as it flows by. Sails acting in this way are aerodynamically stalled. Drag, which is always parallel to the wind, contributes the predominant driving force. The other way sails propel the craft occurs when the craft is traveling across or into the wind. The sails acting as airfoils propel the craft by redirecting the wind coming in from the side towards the rear. By the law of conservation of momentum, the wind moves the sail as the sail redirects downwash air backwards. Air pressure differences across the sail area result in forces on sails including drag and lift. A component of the lift is the main driving force.

9.3.4 Wind power type: Airborne wind power

Airborne wind power (i.e., kite power; also known as airborne wind energy, AWE; tethered aircraft wind power) involves the use of kites and wind to generate mechanical and/or electrical power. The kites themselves may or may not be inflatable, and if inflated, they may be inflated with air, or a gas lighter than air. Often the kites are similar to those used in recreational kite surfing and parasailing. They are lightweight structures that can dynamically change with wind direction and altitude.

Airborne wind power systems have several advantages over conventional ground-based wind systems. The first advantage is that airborne wind power density steadily increases with altitude. This is important because not only does it mean that more power can be generated, than would be generated with a conventional ground-level turbine, but that increase in power will not require an increase in size of the airborne structure. Often, geographic locations that are not suitable for conventional wind farms at surface level have ample wind speeds at higher altitudes. Secondly, airborne wind power systems are often less complex and require fewer materials. Most ground-based wind turbines require massive structural foundations in order to support the large blades, while also requiring significant transportation tasks to move the systems into place (i.e.,, each blade must be transported separately, and typically requires unconventional transportation methods due to its size). Further, ground-based wind turbines take-up land area and output polluting acoustics. When airborne, these systems take-up minimal land area, and their acoustic pollution has little affect on the ground. Some airborne are highly portable.

> **NOTE:** *Airborne power generation platforms can double as platforms for instrumentation and/or communication.*

There are two primary types of airborne wind power system:

1. **Kite propulsion systems** - Kites propulsion systems are similar to sails in that they capture the wind with a concave surface, except they are airborne.
2. **Airborne wind power (non-propulsion) systems** (airborne aerofoil-powered system) - kites/aircraft that produce mechanical and/or electrical power.
   A. **Ground-gen** - Systems that generate the electric power on the ground.
   B. **Fly-gen** - Systems that generate the electric power in the airborne/flying part.

Due to their significant elevation above ground level (and ground-based wind power systems) airborne machines come into contact with the more consistent and stronger wind resource at altitude. At 2,000 feet wind speeds above 8 m/s are blowing more than 40 percent of the time at most locations in the northern hemisphere. Furthermore, power densities (kW/m2) are on par with the world's most favorable sites for ground-based wind. Although boundary layer winds provide reasonable power densities, the jet stream winds of the troposphere at ~10km, where average power densities soar beyond 20 kW/m2 and the total available resource is measured in thousands of terawatts (TW), hundreds of times higher than world energy demand.

Every airborne wind power system has three parts:

1. A grounded part
2. An airborne/flying part
3. A connection between the two primary parts - a tethering cable. The tether allows the airborne structure's altitude to be adjusted and periodically brought back to the ground. In addition, depending on the design of the system, the tethering cable can also serve a conduit or conductive wire that connects the system's electrical components to surface electrical components and/or the electrical power grid. In order to counteract the kite moving sporadically, sometimes two tethers are placed on opposite ends of the kite/aircraft, which work to control the kite's position and keep it positioned correctly.

There are three surfaces that the airborne wind power system may be tethered (i.e., connected) to:

1. Land
2. Stationary structure (e.g., building)
3. Moving structure (e.g., boat)

9.3.4.1 Kite propulsion systems

Kite propulsion systems use wind power (energy) to propel (or aid in propelling) a transportation platform (e.g., ship). These wind power systems use large foil kites, similar to those used in kitesurfing. Kite propulsion systems are similar to sails, except the sail is airborne. These kites are supported and controlled by lines that lead from the kite to the transportation platform.

9.3.4.2 Airborne wind power (non-propulsion) systems

In ground-gen systems, the electrical power of the system is produced by a generator on the ground. The generator on the ground is powered by the mechanical power generated by the wind moving the airborne aircraft/kite. The aircraft's mechanical motion is transferred by the tether to the ground. Generally, this system consists of a turbine positioned within a kite/aircraft, placed in the atmosphere, with the turbine therein is spinning horizontally (instead of vertically). Although the ground-gen systems can stand alone, it is possible to connect multiple aircraft and tethers to a single ground-based generator to maximize output.

Company: KiteGen [kitegen.com]

In fly-gen systems, the electrical generator is located in the airborne/flying part of the system. In fly-gen systems, the tether contains the conductive, insulated cables that run from the generator in the atmosphere to the ground. Unlike ground-gen systems, fly-gen systems do not use the wind to move the aircraft and generate mechanical power that is transferred to the ground. Instead, airfoils (e.g., propellers) are mounted onto the aircraft/kite and generate electrical power by the wind moving past their position. Fly-gen systems can be designed to use their rotors as propellers to thrust the aircraft into the sky, using the generators onboard in "reverse", as engines. Once at the proper altitude, the rotors stop propelling, and allow the aircraft to flow with the wind like a kite, relying on the aircrafts aerodynamics and the wind (i.e., air lift) to keep it afloat. Once at a sufficient altitude, air flow forces the rotors, causing them to spin in the opposite direction, and generate electricity. The electricity is conducted down the tether and into an electrical network. The rotor's controls can be operated from the ground either manually or with an automated system that can adjust the aircraft's flight path according to wind conditions. When it is time to land the aircraft, the rotors can take over as propellers once again and guide the gliding aircraft to the ground with power (i.e., "under control with engine power").

Companies: Makani [makanipower.com];
Altaeros Energy [altaeros.com]

9.3.4.3 Issues with airborne wind power

There are limitations and/or other issues and concerns with airbone wind power systems. Firstly, areas where these airborne wind power platforms are flying are considered no-fly zones, in order to avoid accidents (i.e., air traffic must be re-routed around them). Secondly, weather conditions can be an issue. Depending the operational weather parameters for the system, inclement weather may preclude its airborne functioning, and require its retraction from the atmosphere. During lightning, hail storms, or tornado-strength winds, airborne wind systems must be landed. Icing on the wings will also affect the aircraft performance. Further, without proper monitoring and adhering to safe operation parameters, these systems could break free from their tethering cable, causing them to crash land or cause air accidents, and/or additional air traffic re-routing.

9.3.5 Wind power type: Magnus power effect

The Magnus effect is the commonly observed effect in which a spinning ball (or cylinder) curves away from its principal flight path. The effect involves a force acting on a spinning body in a moving airstream, which acts perpendicularly to the direction of the airstream. The effect can be used to generate propulsion. A "rotor ship", or Flettner ship, is a type of ship designed to use the Magnus effect for propulsion. Rotor ships typically use rotor sails powered by a motor to take advantage of the effect.

9.4 Environmental impact of wind power

The environmental impact of wind power, when compared to the environmental impact of fossil fuels (and other energy sources), is relatively minor. Environmental issues include, but are not limited to:

1. Ecological disturbances - primarily, bird strikes.
2. Noise emission - Some turbines emit low frequency acoustics, which distress complex organisms. Noise emissions are affected by the systems aerodynamics, the location of the blades upwind or downwind of the tower, and the speed of the rotor. Given that the noise emissions from the blades' trailing edges and tips vary by the 5th power of blade speed, a small increase in tip speed can make a large difference in the noise produced.
3. Fires - are particularly dangerous due to the material composition of the blades, and their rotation.
4. Electromagnetic interference - the design of the rotating element can interfere with radar and other electromagnetic transmitting and receiving systems.

10 Solar power

A.k.a., Sun power, solar electromagnetic power.

The sun's energy is the primary source of energy for all surface phenomena and life on Earth. Electromagnetic power (energy) from the sun comes to Earth in the form of radiation. The term 'radiation' denotes the fact that the energy travels as rays, that is, in relatively straight lines. In general, the terms 'solar power/energy' and 'solar radiation' refer to energy from a star (i.e., a sun). The Sun radiates electromagnetic power (energy) "equally" in all directions, and the Earth receives part of this energy. However, because the earth revolves around itself (and the sun), any point on the earth will only receive sunlight periodically (cyclically as a day/night cycle); thus, solar power is directly effective during sunlight (daylight) hours only (although residual heating effects exist). Atmospheric conditions also affect the amount and quality of light reaching the Earth's surface at any given sunlight time.

NOTE: *Solar power is only available during sufficient sunlight conditions.*

The range of electromagnetic power (energy) emitted by the sun is known as the 'solar spectrum', and lies mainly in three regions: ultraviolet, visible, and infrared (of which there are multiple bands of infrared). Therein, the solar spectrum extends from about 0.29 μm (or 290 nm) in the shorter wavelengths of the ultraviolet region, to over 3.2 μm (3,200 nm) in the far infrared. Small amounts of radio waves are also given off by the sun and other stars.

Electromagnetic energy can be discussed in terms of its 'energy distribution' - the spread of energy over a range of wavelengths. This distribution of energy is also known as the 'spectral distribution'. Therein, the measure of radiation may be quantified in terms of the amount of energy falling per second (measured in Watts) per unit area (in square meters, m^2) in each band of 1 μm (1000nm) wavelength.

It is often said that the sun radiates electromagnetic and thermal energy. However, this is not technically correct, because that which is called "thermal radiation" is in fact 'electromagnetic radiation'. Take note here that 'thermal radiation' is 'electromagnetic radiation' generated by the thermal motion of charged particles in matter (i.e., due to the temperature of matter), and when that electromagnetic radiation from those charged particles comes into contact with matter, it will transfer energy (through heating), and raise its internal temperature. Hence, the term 'thermal radiation' describes the source of the radiation and implies that when that radiation contacts matter after its release/generation, that it will likely heat the matter. All matter with a temperature greater than absolute zero emits thermal [electromagnetic] radiation. When the temperature of a body is greater than absolute zero, inter-atomic collisions cause the kinetic energy of the atoms or molecules to change. This results in charge-acceleration and/or dipole oscillation, which produces electromagnetic radiation (i.e., thermal radiation"). The spectrum distribution of EM radiation will reflect the spectrum of accelerations and oscillations that occurred to release that radiation. Herein, the "emissivity" of the surface of a material is its effectiveness in emitting energy as thermal radiation.

Solar [electromagnetic/thermal] radiation heats the earth during the day, while at night the earth re-radiates some energy via heat back into space. In other words, the sun generates solar power (energy), which transfers energy to the earth, heating it and raising its atmospheric and surface temperatures. Whereupon the earth emits electromagnetic/thermal radiation back into space due to its earlier temperature rise and the internal accelerations and oscillations that occurred therefrom.

NOTE: *When light falls on a surface, it can either be reflected, transmitted, absorbed, or varying degrees of all three.*

There are four ways in which solar power (as electromagnetic effects and thermal effects) can be utilized as an energy source. Three of the ways produce electric power, and the fourth produces heating ("thermal power"):

1. Photoelectric effect/power - Solar electromagnetic power can be converted into electrical power through the photoelectric effect, of which there are a variety of sub-types of this effect. This is an active form of solar power.
2. Non-photoelectric effect - Solar electromagnetic power can be directly converted into electrical power through processes other than the photoelectric effect. This is an active form of solar power.
3. Solar heating can be indirectly converted into electrical power (indirect solar-thermal electric). This is an active form of solar power.
4. Solar heating can be used for its direct heating effect (e.g., direct heating, drying, distillation). Whereas this is a form of passive solar power, the other three are active forms of solar power.

These effects/technologies can be combined to increase the overall effectiveness and/or efficiency of useful power generation.

These uses of solar power (energy) can be divided into two categories representing two different types of solar power conversion (energy transfer):

- Active - conversion of the sun's radiated power (energy) to other useful carriers (e.g., electric power and hotter water).
- Passive (passive solar) - the direct use of the sun's

thermal [electromagnetic] power for [passive] heating (of architecture and liquids).

10.1 The solar radiation supply

NOTE: *The available solar energy increases with altitude due to lower atmospheric absorption.*

Irradiance is a measurement of solar power and is defined as the rate at which solar power is falling onto a given surface. In other words, solar irradiance is a measure of how much solar power is contacting a specific location. The unit of power is the Watt (W). Solar irradiance is usually measured as power per unit area, so irradiance is typically expressed as W/m², or for larger amounts, kW/m². The irradiance falling on a surface can and does vary from moment to moment, which is why it is important to remember that irradiance is a measure of power - the rate that energy is transferring, not the total amount of energy.

"Total solar irradiance" is defined as the amount of radiant energy emitted by the Sun over all wavelengths, not just visible light, "falling" each second on a 1 square metre perpendicular plane outside Earth's atmosphere at a given distance from the Sun. It is roughly constant, fluctuating by only a few parts per thousand from day to day. On the outer surface of the Earth's atmosphere the irradiance is known as the 'solar constant', and it is equal to ~1367 Watts per square meter. Solar irradiance at sea level on the equator at noon on a sunny day is ~1000 W/m².

The solar irradiance integrated over time is called solar irradiation, solar exposure, or insolation. However, insolation is often used interchangeably with irradiance in practice. **Solar irradiation** is a measure of the total amount of solar energy accumulated on an area over a period of time. Solar irradiation is expressed as a number of watt-hours per square metre (Wh/m²), and for larger amounts, kWh/m².

CLARIFICATION: *The terms 'irradiance', 'irradiation', and 'insolation' are often used interchangeably to mean the same thing.*

Insolation is another term used to refer to the amount of solar irradiance received over time. The amount of solar energy that strikes a given area over a specific period of time varies with latitude and with the seasons, as well as the weather, and is known as the insolation (incident solar radiation). In other words, the total amount of solar energy that falls over a given time in a given location is called insolation. Whereas, insolation is a measure of solar energy, irradiance is a measure of solar power. Insolation is the power of the sun added up over some time period. For example, if the sun shines at a constant power of 1000 W/m² for one hour, then it has delivered 1 kWh/m² of energy. The amount of power is the product of the power (1000 W/m²) times the length of time (1 hour) and the unit of energy is the kWh. Insolation (measured in kWh) is not the same as power (measured in kW) in the same way that kilometers per hour is not the same as kilometers.

Hence, the power and energy measurements for solar electromagnetic radiation are:

- Solar irradiance (power) - solar radiation every second is expressed as W/m² or kW/m².
- Solar irradiation - solar irradiance over time greater than a second is expressed as Wh/m² or kWh/m².
- Solar insolation - same meaning as solar irradiation.

When the Sun is directly overhead the irradiance (insolation), that is the incident energy arriving on a surface on the ground perpendicular to the Sun's rays, is typically 1000 Watts per square metre. This is due to the absorption of the Sun's energy by the Earth's atmosphere, which dissipates about 25% to 30% of the radiant energy. Irradiance increases with altitude since the radiation passes through less air mass, hence the energy absorption by the atmosphere is less. The amount of solar irradiation received at any particular location on or above the earth's surface varies due in part to atmospheric attenuation ("loss"). This loss of light is caused by contact with air molecules, water vapor, and dust absorbing and otherwise scattering the light. Some of the light is absorbed by atmospheric composition, including but not limited to ozone, water vapor, and carbon dioxide. The amount of light absorbed and scattered depends on atmospheric composition and thickness. The minimum amount of atmosphere the sun's rays have to go through is the condition in which the sun is directly above a given point on or above the earth. This is referred to as the sun's zenith.

Astronomical data is required to calculate how solar insolation varies with time and with the position of the solar device on the Earth's surface. Atmospheric an meteorological conditions impact how much solar radiation will contact a solar powered device. And, the configuration and operation of the solar system itself affect will also impact the ability to interface with the available solar irradiance.

NOTE: *The term* **peak sun hours***, refers to the energy received during total sunlight hours as defined by the equivalent number of hours it would take to reach that total energy value had solar irradiance averaged 1000 W/m2. In other words, peak sun hours are the number of hours per day that solar insolation equals 1000 W/m². This term is interchangeable with kWh/m² /day.*

Solar power systems operate within a set of three types of parameters: astronomical parameters; atmospheric and meteorological parameters; and [solar power] device parameters. The maximum transfer of energy between the sun and a solar power system occurs when parameters within these three sets are optimized. The set parameters under which a solar power system can produce the maximum power is called 'maximum power' and/or 'optimum power transfer'. For instance, the voltage at which photovoltaic module can produce maximum power is called 'maximum power point' (or peak power voltage). Maximum power varies with solar radiation, ambient temperature, and solar cell temperature.

10.2 Astronomical parameters

The sun changes its position relative to any point on the earth from morning to night, and from one season to another. The following are astronomical parameters relevant to the positioning and orientation of the contact interface of a solar powered device.

1. **Relative solar positioning** - The position of the Sun in the "sky" relative to a point on earth is defined by its:
   - Altitude angle 'α' (solar elevation angle)
   - Azimuth angle 'Ψ'
2. **The Earth's orbit** - The Earth orbits the Sun with one revolution per year in an elliptical orbit with the Sun at one of the foci of the ellipse. The orbit's two foci are sufficiently close together such that the orbit is approximately circular (slightly elliptical) -- the distance to the sun from the perihelion (the point in its orbit closest to the sun, ~ January 2) is approximately 3% less than its distance from the aphelion (its furthest distance, ~July 3). The sun moves faster across the horizon at perihelion, than at aphelion. Because the orbit is approximately circular, the effect of the orbit on solar irradiance remains essentially constant throughout the year as the earth orbits the Sun. The actual energy received at any distance from the Sun is determined by the inverse square law. Thus a 3% change in distance gives rise to a 6% change in the irradiance.
3. **The Earth's rotation** - The earth rotates about its inertial plane, which passes through the north and south "poles" of the planet. The earth rotates about this axis once each "day" (approximately 24 hours). The earth's rotation of once per day defines this planet's day and night cycle. As the Earth rotates the insolation at any point on its surface rises to a maximum at mid day and falls to zero during the night as the earth presents a different region ("face") toward the sun. For maximum efficiency the orientation of a solar should follow the Sun as it passes through the "sky".
4. **Latitude** - A solar device placed on the ground will only receive the maximum insolation when the sun is directly overhead. Because the Earth is roughly spherical, the angle between the plane of the earth's surface and the incident solar radiation will gradually increase from 90° as the device is repositioned away from the equator toward the upper and lower latitudes by an angle 'Θ', equal to the latitude of the device's location on earth. At this point the altitude angle 'α' of the Sun will be (90 - Θ) degrees. Because of the increased inclination of the earth's surface the insolation received by a device placed on the surface will gradually decrease. This shift in the position of the sun relative to a point on earth can be overcome by inclining the interface plane of the device so that it is perpendicular to the incoming solar rays'. The amount of elevation from the horizontal, the tilt angle, should be equal to the latitude angle 'Θ' of the location of the device. For maximum interface, the axis of inclination should be relatively perpendicular to the polar axis.
   - Note that the polar axis ("true north" and "true south") is not the same as the compass bearing of north and south, because the magnetic poles do not necessarily line up exactly with the geometric poles. In other words, true north/south is not the same as magnetic north/south. The angle between the magnetic and geographical meridians at any place is called the magnetic declination or variation, which can be as much as 20 degrees or more in deviation. It is expressed in degrees east or west to indicate the direction of magnetic north/south from true north/south. If a compass is being used to orient a solar device, then the difference must be corrected for, which varies from place to place.
5. **The Earth's tilt** - The earth's rotational axis is tilted ~23.45 degrees from the plane of its orbit. This tilt is essentially constant, maintained in that direction due to the gyroscopic action of the earth's rotation, and always points in the same direction relative to the stars, so that the north pole points towards the star polaris, the north star. Over very long time periods however, measured in thousands of years, the direction of earth's axis slowly changes due to gyroscopic precession. The fixed orientation in space of the earth's axis as it orbits the Sun determines the length of the day and creates the world's seasons. As a result of the Earth's tilt, the intensity of the insolation varies during the year giving rise to the seasons. This is not because tilt causes a point on the earth's surface to move closer to or further from the sun. The change in distance is negligible. It is because of three factors, which together reduce both the intensity and daily duration of the insolation during "winter" months:
   - The earth's tilt changes the angle of incidence of the solar radiation, changing its insolation per unit area. The intensity of solar radiation is

largely a function of the angle of incidence, the angle at which the Sun's rays strike the Earth's surface.
- The tilt also changes the path length of the radiation through the atmosphere, which in turn changes the amount of the Sun's energy absorbed by the atmosphere.
- The tilt changes the number of sunlight hours.
- The declination in the elevation of the Sun varies during the course of the year between minus 23.45° in the summer and plus 23.45° in the winter. The angular position of the Sun at its highest point in the sky with respect to an observation point on the plane of the equator and is called the solar declination 'δ' (not to be confused with magnetic variation, also called the declination). Accounting for the solar declination, the altitude angle α of the sun is (90 - Θ ± δ) degrees. Hence, the inclination angle of a solar interface from the horizontal (for maximum efficiency) should therefore be (Θ ± δ) degrees, and the device should be able to follow this variation in declination throughout the year.

6. **Time** - Sunlight is only available during sunlight hours.

10.2.1 Atmospheric and meteorological parameters

During sunlight hours the magnitude and quality of solar radiation contacting the earth's surface is dependent upon atmospheric/meteorological conditions. These conditions affect not only solar radiation, but may affect the performance of the solar power system itself.

1. **Meteorological presence (weather)** - will significantly determine the amount of solar radiation available to a solar device. Cloud cover, dust, precipitation, and air pollution, as meteorological conditions, will impact air composition/density and interface clarity, and hence, the effective power output of a solar device.
2. **Air density (atmospheric density)** - the mass per unit volume of atmosphere. The greater the mass, the more likely absorption and scattering become.
3. **Air pressure (atmospheric/barometric pressure)** - the pressure exerted by the weight of air in the atmosphere of Earth as a meteorological condition.
4. **Temperature (atmospheric temperature)** - thermal quality of atmosphere.
5. **Humidity (atmospheric humidity)** - the amount of water vapor in the air.

NOTE: *In general, solar systems, particularly thermal, work in relatively cold weather, because the device interfaces with the Sun's radiation. However, ambient air temperature, and the design of the solar device, will impact functioning.*

10.2.2 Solar power system parameters

These are parameters specific to the solar powered system itself. The orientation of the interface of the solar device/collector/array with respect to the position of the Sun is a major determinant in the efficiency of a solar power system. The amount of energy transferred to/through (i.e., "captured by") a solar system can be maximised if the collector can follow the ecliptic path of the sun so that the plane of the collector or array is always perpendicular to the direction of the sun. In order to get the most power output from a solar device, it needs to point in the direction that captures the largest quantity of solar rays. In other words, the amount of irradiance contacting a collector or array is directly proportional to the area of the radiation wave-front it intercepts. For optimum energy capture the collector must be perpendicular to the Sun's rays (i.e., when the angle of incidence is 90°). For a flat plate on horizontal ground this occurs only when the Sun is directly overhead. Unless the solar device is located within the equatorial region of the Earth, a 90° angle of incidence is not possible due to the position of the sun relative to the system's position on Earth.

When the incident energy is not perpendicular to the collector, the angle of incidence is (90° - Θ) and the effective area of the collector is AcosΘ, where A is the area of the collector, and Θ is the deviation from perpendicular of the radiation.

If the Sun's radiation is not perpendicular to the Earth, the transit path through the Earth's atmosphere will be longer, and hence, the energy absorbed on the way to the collector or array will be greater, because it will encounter more air (as in, air mass).

NOTE: *The 'air mass' is a dimensionless quantity defined as the ratio between the actual path length of the solar radiation through the atmosphere and the vertical path length through the atmosphere at sea level.*

The effect of a longer route through the atmosphere is to increase the energy absorption (or lost energy) by a factor of 1/cosΦ, where Φ is the deviation from perpendicular of the radiation, also called the zenith angle. Thus, in the polar regions as Φ approaches 90 degrees (cosΦ>0), the insolation is very low, even if the collector is pointed directly at the sun, due to the longer path through the atmosphere.

In concern to a solar device's mounting orientation toward the sun, there are two types of orientation:

1. Fixed [static] mount (fixed tilt) - do not track the sun.

2. Automatic tracking [passive and active] - tracks the sun.

Note that a fixed tilt design can be adjusted (re-mounted) at least twice a year to give a meaningful increase in power output. There are two types of automatic tracking: passive and active (both are mechanical). Automatic mechanical tracking systems make it possible to track both the azimuth and the elevation of the sun's position to maximise energy capture.

Automatic tracking systems include:

1. **Azimuth tracking:** Azimuth tracking keeps the device's interface pointing at the sun as the earth rotates.
    A. **Passive systems** provide the simplest form of azimuth tracking. They have no motors, controllers or gears, and they don't use up any of the energy captured by the collector. They depend on the differential heating of two interconnected tubes of gaseous refrigerants, one on either side of the collector. If the collector is not pointing towards the sun, one side heats up more than the other and vaporises its refrigerant. The resulting change in weight is used in a mechanical drive mechanism to turn the collector towards the Sun where it will remain when the temperature and weight of the two tubes will be balanced.
    B. **Active tracking** is also possible by employing temperature sensors and a control system with linear actuating motors taking their drive power from the system.
2. **Altitude/elevation tracking:** Elevation tracking enables the interface to follow the seasonal variations in the sun's altitude, but alignment is less accurate than for azimuth tracking. Compared with the daily variations in insolation, the seasonal variations are slow and the range of the variation, due to the solar declination is much more restricted. Because of this, reasonable efficiency gains can be obtained simply by manually adjusting the elevation of a static mount every two months.
3. **Dual axis tracking:** Combining azimuth and elevation tracking enables the interface to capture the maximum energy using the smallest possible collectors.

10.2.2.1 Maximum power point tracking (MPPT)

Maximum Power Point Tracking (MPPT) is algorithm that included in charge controllers used for extracting maximum available power from power conversion module under certain conditions. Jacobi's Law states that a power source will deliver its maximum power to a load when the load has the same impedance as the internal impedance of the power source. Note that batteries are not generally designed for interconnection with power conversion systems, and the mismatch results in efficiency losses. Thus, a power point tracker is a form of voltage regulator that is placed between the power system and a battery. It presents an ideal load to the power system allowing it to operate at its optimum voltage to maintain its full power (wattage), regardless of the battery voltage. The voltage at which power module can produce maximum power is called 'maximum power point' (or peak power voltage).A variable DC/DC converter in the module automatically adjusts the DC output from the module to match the battery voltage. It is not enough however to match the voltage at the specified maximum power point (MPP) of the PV array to the varying battery voltage as the battery charges up. Due to changes in the intensity of the radiation falling on the array during the day as well as to changes in the ambient temperature, the operating characteristic of the PV array is constantly changing and with it the MPP of the PV also changes. For optimum power transfer, the system needs to track the MPP as the solar intensity and ambient temperature changes in order to provide a dynamic reference point to the voltage regulator.

10.2.3 Solar power interface types

The type of solar device will determine the type of interface. In general, either:

1. The material composition of the solar device itself acts as the interface (e.g., photovoltaic cells/panels or direct solar collector), or
2. A mirror-receiver system is used. In a mirror-receiver system, the mirrors concentrate/reflect the solar radiation onto a receiver that absorbs the radiation; e.g., solar concentrator).

A solar collector is a heat collecting surface that intercepts the sun's radiated energy and heats up a thermal working fluid. In practical thermal systems it is usually more convenient to focus the suns thermal energy onto a small receiver in order to obtain a higher temperature rise of the working fluid. Such collectors are called concentrators. Typically, concentrators are constructed from parabolic mirrors, which reflect the sun's parallel rays on to a single spot at the focus of the mirror.

> **UNITS NOTATION:** *Note that the units used to express the degree of concentration of the mirror system are similar to the magnification factor of a lens, and are called* **suns** *(this unit is not a precisely defined quantity).*

There are several types of concentrators:

1. **Parabolic dish** - a shaped mirror that redirects solar energy onto a suitable heat absorber/receiver

located at the focus.

2. **Parabolic trough** - a shaped mirrors that redirects solar energy onto a pipe-like heat absorber located running through the focus. Generally, this setup forms a type of "solar furnace" used to raise steam to drive a turbine generator.
3. **Power tower** - uses a large array of parabolic mirrors focused on a solar furnace mounted on the top of a tower. Because of the long focal length, the mirrors are almost flat. Generally, this setup forms a type of "solar furnace" used to raise steam to drive a turbine generator.
4. **Heliostats** - sun tracking mirrors used to reflect solar radiation in a fixed direction (such as toward a solar panel or thermal receiver).

NOTE: *For most of these systems, the amount of energy captured, and hence, the temperature rise of the absorber will be proportional to the area of the dish-like mirror.*

10.2.4 Solar power system monitoring

Solar devices require a "clear" (i.e., unobstructed) interface with the solar electromagnetic radiation to function optimally. If the device is covered with a material (e.g., dust, snow, debris), then the device will produce less power relative to the obstructing material type, and amount of material, covering the device. Hence, solar devices require periodic cleaning to ensure the device's interface surface is free of extraneous material.

Note that the solar interface itself may degrade with exposure to sunlight, and may require periodic replacement.

10.3 Photoelectric power: Direct transfer of solar electromagnetic energy to electric power

NOTE: *These systems are sometimes referred to as* **photoelectric transducers***, wherein a transducer is a device that transfers ("converts") energy from one carrier ("form") to another carrier ("form"). The majority of these systems transfer electromagnetic power (energy) to mechanical displacement of charges, which produces electrical power. The other systems transfer a non-electrical physical quantity, such as temperature or sound, to an electrical signal.*

At the top level, the photoelectric and photovoltaic effect refers to the direct production of electric power (as direct current (DC) voltage) through the contact of electromagnetic radiation (energy) with a semiconducting material. Note here that the term 'photo' means 'light', which is another term for electromagnetic radiation. The photo-electric/-voltaic effect (or photoemission) refers to the observed movement of free charge carriers (or electrons) when light (electromagnetic radiation) contacts a material. Electrons "emitted" in this manner may be called "photoelectrons". The photoelectric effect is an electromagnetic, physical and chemical phenomenon.

The photoelectric and photovoltaic effects are closely related, and the terms are sometimes used interchangeably. In either case, light is absorbed, causing excitation of an electron or other charge carrier to a higher-energy state. In other words, "photons" (dielectric perturbations) carry the potential of moving electrons out of materials. Or, said another way, it is the direct conversion of light into electrical power (voltage and direct current) at the atomic level. In either case, an electric potential (or voltage) is produced by the separation of charges, and the light has to have a sufficient energy to overcome the potential barrier for excitation.

CLARIFICATION: *In general, the 'photoelectric effect' is the physical phenomenon responsible for the creation of an electrical potential difference (voltage) in a material when exposed to light. However, when speaking about specific occurrences of the photoelectric effect, the terms photoelectric and photovoltaic mean something different. In specific, the ways in which the electrons are emitted in the photoelectric effect and photovoltaic effect create the difference between them. They differ in definition as the steps of progression are different in each case. The main scientific difference between the two processes is that in the photoelectric effect, the electrons are emitted into a vacuum space (usually via ballistic conduction), whereas in the photovoltaic effect, the emitted electrons directly enter a new material (usually separation is via diffusion).*

There are several basic types of photoelectric device (i.e., photoelectric cell), corresponding to the different forms of the photoelectric effect that they employ. A **photoelectric cell** is defined as a device whose electrical characteristics, such as current, voltage, or resistance, vary when exposed to light.

Modules (a number of connected cells), are connected/arranged into arrays (connected modules/cells). The power production module/cell has no moving parts. All forms of the effect utilize semiconductor material in order to generate the effect. Solar electromagnetic energy may be directly converted to electric potential energy via the following effects/methods. Note that any devices employing these effects/methods could be referred to as a solar cell, although photovoltaic cells are most commonly referred to as solar cells:

1. **Photoconductive effect (photoconductive cell/ photoresistor)** - light "frees" electrons from their valence bonds in a semiconductor material, while reducing the materials resistance (increasing electrical conductivity). The photoconductive cell

is a two terminal semiconductor device whose terminal resistance will vary (linearly) with the intensity of the incident light. For obvious reasons, it is frequently called a photoresistive device.

2. **Photoemissive effect (photoemissive cell, photocell, electric eye)** - light knocks electrons from a cathode to an anode, making a current flow through an external circuit.
3. **Photoelectric effect (photoelectric cell) -** light makes electrons move out of semi-conducting material and into a vacuum space.
4. **Photovoltaic effect (PV cell, photovoltaic cell, or solar cell)** - light makes electrons move between layers of semiconducting material, producing a voltage and a current in an external circuit. A photovoltaic cell may also be used as a photodetector (e.g., infrared detectors), detecting light or other electromagnetic radiation near the visible range, or measuring light intensity. A solid state electrical device that converts the energy of light directly into electricity by the photovoltaic/ photoelectric effect. Photovoltaic cells (PV cells) are the building blocks of photovoltaic modules, otherwise known as solar panels. The operation of a photovoltaic (PV) cell requires 3 attributes:
   - The absorption of light, generating either electron-hole pairs or excitons.
   - The separation of charge carriers of opposite types.
   - The separate extraction of those carriers to an external circuit.
5. A **photo-electrochemical cell (photo-electrolysis cell)** - may either be a type of photovoltaic cell (like that developed by Edmond Becquerel and modern dye-sensitized solar cells), or a type of cell that splits water directly into hydrogen and oxygen using only solar illumination. Photo-electrochemical cells or PECs are solar cells that produce electrical energy or hydrogen in a process similar to the electrolysis of water. Photoelectrolysis is the process of using sunlight directly to decompose water into hydrogen and oxygen through the use of semiconductor material similar to that used in photovoltaics. In other words, by passing light (electromagnetic radiation) through water it is possible to split the water into its component parts.
6. **Photogalvanic effect (photogalvanic cell)** - a special case of the so-called Becquerel effect, in which the influence of light on the electrode potential is due to a photochemical process in the body of the electrolyte (as distinct from photochemical or photoelectric processes in the surface layer of the electrode, which are the basis of the original Becquerel effect).
7. **Photomagnetic effect (photomagnetic cell)** - a material acquires (and in some cases loses) its ferromagnetic properties in response to light. The current model for this phenomenon is a light induced electron transfer, accompanied by the reversal of the spin direction of an electron. This leads to an increase in spin concentration, causing the magnetic transition. (Mahmoud, 2015)

10.3.1 Photovoltaic cells

The most common effect used to create solar panels is the photovoltaic effect. Here, solar power converts sunlight into DC voltage electrical power using photovoltaics. Solar cells, also called photovoltaic (PV) cells get their name from the process of converting light (photons) to electric current (voltage), via a mechanistic effect known as the PV effect. Photovoltaic (PV) technologies may be divided into three generational categories:

1. **Wafer-based PV** (also called 1st generation PV, crystalline silicon, and bulk PV) - This first generation PV cell technology monocrystalline silicone (monosilicon) cells and polycrystalline silicone (polysilicone) cells.
2. **Thin-film cell PV** (also called 2nd generation PV or thin film solar cell, TFSC) - This second generation PV cell technology includes amorphous, protocrystalline, and nanocrystalline cells.
3. **Organic photovoltaic cell** (OPVC) - This third generation PV cell technology uses organic, electronic conductive polymers or small molecules for light absorption and electrical charge transport.

All photovoltaic cells utilize a semiconductor P-N junction. Photovoltaic cells contain p-type materials flush with n-type materials. Sunlight provides the energy to make the current flow from the n-type to the p-type.

In general, the photovoltaic process occurs as follows:

1. "Photons" in rays of sunlight contact a solar panel and are absorbed by semiconducting material.
2. Electrons in the semiconducting material are excited from their current molecular/atomic orbital.
3. Once excited an electron can either dissipate the energy as heat and return to its orbital, or travel through the cell until it reaches an electrode. Once an excited electron reaches an electrode a current is created. The chemical bonds of the material are vital for this process to work.
4. An array of solar cells converts solar energy into a usable amount of direct current (DC) electricity.
5. An inverter can convert the power to alternating current (AC).

Photovoltaic diodes (a.k.a., photodiodes, PV cells, or solar cells) generate an electric current when light of sufficient magnitude impinges on a semiconductor lattice connected to a P-N junction. If the photon energy in the light energy is less than the band gap (energy range in a solid where no electron states can exist), the energy is simply dissipated as heat, and no electrons are released into the conduction band and no current flows. However, if the energy level of the photons is equal to, or higher than, the band gap of the semiconductor material, it will cause the covalent bonds in the semiconductor to be "broken" as electrons jump the band gap into the conduction band. Both the electron and the vacant site left behind by the electron in the valence band (the hole) then act as free charge carriers and contribute to the possible current. Once a photon has caused the release of an electron, any photon energy it had in excess of the band gap energy will be dissipated in the form of heat. Photons pass through the crystal lattice until they are absorbed as heat or until they give up their energy by causing the generation of electron hole pairs and the release of an electron across the band gap.

Photovoltaic cells can be stand-alone systems, or incorporated into other useful materials. Hence, the structural design of photovoltaic cell/panel can take several forms:

- Flexible panels - a flexible stand-alone PV cell.
- Rigid panels - a rigid stand-alone PV cell.
- Shingle and tile panels - a tile (e.g., roof tile) could be a PV cell.
- Window panels - a window could be a PV cell.

10.3.2 Device specifics

Specific devices will be accompanied by the following referential data:

1. Manufacturer - Organization/company name; brand
2. ID - Specific solar module identification code; module name
3. Rating - standard testing conditions; nameplate rating under laboratory conditions
4. Efficiency (%) - Output per input light irradiance using STC; energy conversion efficiency; module efficiency
5. Tier - Solar panel efficiency Trier 1 is highest, 5 is lowest.

10.4 Solar non-photoelectric power: Direct transfer of solar electromagnetic energy to electric energy

Solar radiation can be directly converted into electrical power without the use of the photoelectric effect, via five methods:

1. **Thermoelectric or thermophotovoltaic effect**- The conversion of temperature differences to electric voltage and vice versa. Thermophotovoltaic (TPV) cells involve a direct energy transfer through heat [mode] to electricity via photons (as in, generating electricity from infra-red radiation). A thermoelectric device creates voltage when there is a different temperature on each side. Conversely, when a voltage is applied to it, it creates a temperature difference.
2. **Thermionic emission effect** - The process by which free electrons are emitted from the surface of a metal when external heat energy is applied. Heat energy transfer induces the flow of charge carriers from a surface or over a potential-energy barrier. This occurs because the thermal energy given to the carrier overcomes the work function of the material. Thermionic emission occurs in metals that are heated to a very high temperature. In other words, thermionic emission occurs, when a large amount of external energy transferred via heat is supplied to the free electrons in the metals (i.e., raises the internal energy of the substance and causing the valence electrons to gain enough energy to break their bonding with the parent atom, whereupon they become "free" (i.e., acquire kinetic energy).
3. **Ferroelectric effect (ferroelectricity)** - Ferroelectricity is a property of certain materials that have a spontaneous electric polarization that can be reversed by the application of an external electric field. Certain solids exhibit spontaneous electric polarisation when exposed to light (Read: light has an electric field, as in electromagnetic field). The nonlinear nature of ferroelectric materials can be used to make capacitors with tunable capacitance.
4. **Magnetohydrodynamic effect** (MHD; also magneto fluid dynamics or hydromagnetics) - The study of the magnetic properties of electrically conducting fluids. Examples of such magneto-fluids include plasmas, liquid metals, and salt water or electrolytes.
5. **Electro-gas-dynamic effect** - The process of creating electrical power by converting the kinetic energy contained in a flowing, high-pressure, ionized combustion gas.

10.5 Solar heating electric power: Indirect transfer of solar energy to electric energy

The production of electric power by solar heating (energy transfer via heat) requires two stages. First, solar radiation (thermal energy) heats a working fluid, which is then used in a second energy transfer stage to generate the electric power. A solar "thermal" power system usually involves an array of mirrors to concentrate the sunlight on to an absorber. The absorbed energy is then used to power a heat engine (and turbine), which in turn drives a rotary generator. In large scale systems, the heat engine includes a turbine driven by steam or other vaporous working fluid. The steam (or other vapor) is produced by the concentrated heating effect in the heat engine, which transforms the working fluid therein.

In other words, electric power via solar heating uses concentrated light from the sun to heat a working fluid in a heat engine (and turbine), which turns a generator to make electricity (AC voltage internally). The working fluid that is heated by the concentrated sunlight can be a liquid or a gas. Different working fluids include water, oil, salts, air, nitrogen, helium, etc. In small scale systems the heat engine may be a Stirling engine. Other engine types include steam engines, gas turbines, etc.

Hence, solar power can be indirectly/secondarily converted into electrical power through the thermodynamic process. The thermodynamic process occurs when energy in solar radiation is transferred via heat into shaft work/power through a heat engine (via the Rankine cycle, Stirling cycle, or Brayton cycle), and then, shaft work (mechanical power) is converted into electrical power through a generator (e.g., alternator).

Herein, devices can be constructed to reflect and absorb/collect solar energy. For instance, a **solar thermal collector** supplies heat by absorbing sunlight for the purpose of either direct heating or indirect electrical power generation from heat. Similarly, a **concentrating solar power (CSP) system**, concentrates the sun's radiated energy using reflective devices such as troughs or mirror panels to produce a focused super heated thermal power source that is then used to generate electricity.

Summarily, a solar "thermal" power system mainly consists of a: solar energy collector field, a fluid flow distribution system, a suitable working fluid, a heat engine (and turbine), an electric generator, and a control system. Amongst the many available systems, the two most generic and common are: the central receiver thermal electric power system and the distributed solar thermal electric power system. These two systems have a comparatively high efficiency. In the central receiver concept large arrays of sun-tracking mirrors known as heliostats reflect the solar flux on to the central receiver boiler at the top of the tower. Here concentration ratios power a turbine (steam type) that in turn powers a generator.

10.6 Solar heating (passive): Direct thermal heating

Solar radiation can be used and concentrated for its heating ("thermal") properties. Uses of direct solar "thermal" energy include, but are not limited to:

1. **Solar water heating** - Solar water heating systems, which contain a black solar collector that faces the sun, and either heats water directly or heats a "working fluid" that, in turn, is used to heat water.
2. **Solar walls** - Transpired solar collectors, or "solar walls," which use solar energy to preheat ventilation air for a building.
3. **Solar evaporation** - Solar desalination and solar distillation has been in practice for a long time as a means for evaporating water.
4. **Solar drying** - Solar drying, dehydrating, evaporation, desiccating, and cooking has been in practice for a long time (e.g., food dehydration; salt evaporation ponds; clothes drying; solar oven/ cooker).
5. **Solar ponds** - Solar ponds are used for the collection and storage of thermal energy (as internal energy). The energy can be transferred from the solar ponds for any suitable heating ("thermal") application. Therein, solar pond systems may be used for thermal energy storage, desalination, and electricity generation. A solar pond is a large scale solar collector with an integrated arrangement for storage of heated water. A solar pond is a pool of water (generally, salt water) that collects and stores solar thermal energy (as internal energy). The saltwater naturally forms a vertical salinity gradient also known as a "halocline", in which low-salinity water floats on top of high-salinity water. The layers of salt solutions increase in concentration (and therefore density) with depth. Below a certain depth, the solution has a uniformly high salt concentration. When the sun's rays contact the bottom of a shallow pool, they heat the water adjacent to the bottom. When water at the bottom of the pool is heated, it becomes less dense than the cooler water above it, and convection begins. Solar ponds heat water by impeding this convection. Salt is added to the water until the lower layers of water become completely saturated. High-salinity water at the bottom of the pond does not mix readily with the low-salinity water above it, so when the bottom layer of water is heated, convection occurs separately in the bottom and top layers, with only mild mixing between the two. This hot, salty water can then be pumped away for use in electricity generation, through a turbine or as a source of thermal energy. There are ad-/dis-vantages with solar ponds:
   A. The accumulating salt crystals have to be removed and can be both a valuable by-product and a maintenance requirement.

B. No need for a separate collector for this thermal storage system.
C. The extremely large thermal mass means power is generated night and day.
D. Due to evaporation non-saline water is constantly required to maintain salinity gradients.

11 Geothermal power

TERMINOLOGY:

- **Geothermal heat pump:** Devices that utilize the relatively constant temperature of the Earth's interior as a source and sink of heat for both heating and cooling. When cooling, heat is extracted from the space and dissipated into the Earth; when heating, heat is extracted from the Earth and pumped into the space.
- **Geothermal plant/facility:** A combination of devices (forming a facility, or "plant") in which the prime mover is a steam turbine. The turbine is driven either by steam produced from hot water or by natural steam that derives its energy from heat found in rocks or fluids at various depths beneath the surface of the Earth. The energy is extracted by drilling and/or pumping.

Geothermal energy is defined as thermal energy (heat) from the earth. Geo means "earth," and thermal means "heat" in Greek. Heat flows outward from the earth's interior. The earth's heat content is about 10^{31} joules. This heat naturally flows to the surface by conduction at a rate of 44.2 terawatts (TW) and is replenished by radioactive decay at a rate of 30 TW. The earth has been emitting heat for approximately 4.5 billion years, and will continue to emit heat for billions of years into the future.

> **NOTE:** *According to archaeological evidence, geothermal resources have been used by humanity in the form of naturally occurring hot springs for more than 10,000 years.*

Heat emanating from the earth's interior generates magma (a molten and semi-molten rock mixture). This mixture is usually made up of four parts: a hot liquid base, called the melt; minerals crystallized by the melt; solid rocks incorporated into the melt from the surrounding confines; and dissolved gases. Magma exists in the mantle and lower crust, and sometimes bubbles to the surface as lava. However, the crust of the earth traps most of the mantle (and its heat) beneath it. The earth's crust effectively acts as an insulator, which must be pierced by fluid conduits (of magma, water or other) to release the heat underneath. Because magma is less dense than surrounding rock, it rises within the mantle and through conduits within the crust. As the magma rises, it heats the surrounding rock, as well as water contained in rock pores and rock fractures under the earth's crust. Heat is carried to the surface by fluid circulation, either through magma conduits, hydrothermal circulation, hydrocarbon wells, drilled water wells, or a combination of these.

> **NOTE:** *There are many natural geothermal features on the surface of the earth, including but not limited to: geysers; hot*

springs; volcanoes; steam vents; underwater hydrothermal vents; and mud pots. When a volcano erupts, the mantle of the earth flows up through the crust to the surface as lava.

Geothermal energy exists in different carriers ("forms") all over the Earth (by steam vents, lava, geysers, or simply dry heat). Magma, heated rock, and heated water are all sources of geothermal energy. In order for a geothermal power system to operate, the heat must be carried to the surface by fluid circulation, either through magma conduits, hot springs, hydrothermal circulation, oil wells, drilled water wells, or a combination of these. This circulation sometimes exists naturally where the crust is thin: magma conduits bring heat close to the surface, and hot springs bring the heat to the surface. Once at the surface or the system's location, the thermal energy can be captured and used directly for heating, or it can be used as (or converted to) steam, and directed through a turbine-generator to produce electrical power.

NOTE: *Geothermal steam and hot water can reach the surface in two ways: through naturally occurring surface features such as geysers and fumaroles, or through man-made wells that are drilled down into the reservoir to harvest the energy.*

Natural heat from within the Earth may be transferred through a geothermal power system for the production of electric power, space heating, or industrial steam. Geothermal power uses heat from geothermal fuel (i.e., the inner earth) to heat water or another working fluid. The working fluid is then used as space heating, as industrial steam, or to turn a generator, thereby producing electricity.

Unlike other renewable energy sources, geothermal systems are considered "baseload". This means they can continue to operate throughout all seasonal changes on the earth's surface, and are not dependent on changing surface factors, such as the presence of wind or sun. Geothermal electrical power systems are capable of producing energy continuously (i.e., 24/365 - 24 hours a day and 365 days a year, any time and every day).

NOTE: *Rock is considered a reasonably good conductor of heat. Conversely, air is considered a reasonably poor conductor of heat.*

In order to obtain enough energy to generate electricity, geothermal power plants rely on heat that exists a few kilometers below the surface of the Earth. In some areas, the heat can naturally exist underground as pockets steam or hot water. However, most areas need to be "enhanced" with injected water to create steam. The heat energy transferred from the mantle to the waters in the earth's crust can be harnessed to create electricity.

Geothermal power stations are similar to other steam turbine thermal power stations – heat from a fuel source (in the case of geothermal, the earth's core) is used to heat water or another working fluid. The working fluid is then used to turn a turbine connected to a generator, thereby producing electrical current. The fluid is then cooled and returned to the heat source.

11.1 Geothermal sources

NOTE: *Water boils underground and generates steam at temps of 165˚C and pressure of about 100 psi.*

The type of geothermal field/reservoir will determine the type of geothermal power system (plant or station) that can be built on the geothermal site. There are two primary carriers of geothermal energy, and hence two primary geothermal field types:

1. Water-based fields - including steam and hot water.
2. Hot and molten rock-based fields - including hot and molten rocks, but no water.

Therein, there are four common types of geothermal fields/reservoirs; three of which are water-based and the fourth is rock-based:

1. **Hot water fields/reservoirs** are geothermal aquifers that contain reservoirs of water between ~60°-100° C. In general, these fields are used for direct thermal heating or as part of a binary power plant.
2. **Wet steam fields/reservoirs** are geothermal reservoirs that contain pressurized water that is above boiling point (100° C), so that when the water is pumped to the surface it becomes steam.
3. **Dry steam fields/reservoirs** are geothermal reservoirs that contain pressurized and "superheated" water that is significantly above boiling point. The "superheated" water/steam is pumped to the surface for use. Dry steam fields are rarer than wet steam fields.
4. **Hot and molten rock fields** are geothermal locations with hot rocks.

Electricity production from each type depends on field temperatures, and when water is present, pressures also.

NOTE: *Most geothermal reservoirs are found deeper underground than groundwater reservoirs.*

11.2 Geothermal power Types

Geothermal power can be harvested in a number of different ways. Firstly, water can be used via steam to power a turbine and produce electrical power, or the hot water itself can be used for heating purposes. Secondarily, heated rocks can be used to power a closed-loop heat pump with an internal turbine system to produce electricity, or the temperature difference between an underground area and the surface can be used as part of a heating and cooling system.

There are five water-based geothermal power systems include:

1. **Dry steam power (plant/station)** - The direct use of natural underground sources of steam of 150°C or greater to power turbines. In this case, the underground source of steam (water) is from a dry-steam reservoir. The steam is piped directly into the turbine, which is connected to a generator that generates electrical power. Steam directly from the geothermal reservoir runs the turbines that power the generator, and no separation ("flashing") is necessary, because the geothermal wells only produce steam. Dry steam systems use water in the earth's crust, which is heated by the mantle and released through vents in the form of steam. Dry steam geothermal systems have pipes that are drilled into the site and used to trap and redirect the steam. The steam is then used to turn turbines connected to a generator to produce electricity. The underground water reservoirs that feed such a system is refilled when rain falls on the land. The rainwater eventually soaks back into the crust of the earth. Sometimes, however, dry steam reservoirs do not refill themselves in a very consistent manner.
2. **Wet steam power (a.k.a., flash steam plant/ station)** - This system is similar to a dry-steam power system, but water is recycled back into the thermal well. In this case, the underground source of steam (water) is a wet-steam reservoir. This systems use a well that is drilled into the geothermal site to release the steam. After the steam is piped up, and passes through the geothermal power system, it is sent into a condenser, which cools it. This cooled water is then pumped back into the well (i.e., recycled). The water is heated again by the geothermal source, and released as steam, and the process repeats. This process can happen over and over with a minimal loss of water. Wet steam power system use naturally occurring sources of underground hot water and steam, which is "flashed" into steam, used to power a turbine-generator, and then condensed and recycled. Hot water and steam are pumped into a low-pressure area on the surface known as a flash tank. The water evaporates rapidly (i.e., "flashes") into steam, which is funnelled through a turbine connected to a generator to produce electrical power. Any remaining water can be flashed in a separate tank to extract more energy. The water is then condensed and recycled back into the earth. Flash steam systems pull deep, high-pressure hot water into lower-pressure tanks and use the resulting flashed steam to drive turbines. They require fluid temperatures of at least 180°C (usually higher). This is the most common type of geothermal station in operation today. The hot water flows up through wells in the ground under its own pressure. As it flows upward, the pressure decreases and some of the hot water boils into steam.

   NOTE: *A 'condenser' is a heat exchanger, which condenses a substance from its gaseous to its liquid state.*

3. **Binary cycle power (plant/station)** - Water heated underground (at ~107°-182° C) is piped to the surface where it is passed through a heat exchanger, containing a pipe with a secondary working fluids (a fluid with a lower boiling point). After being heated by the water piped up from underground, the working fluid is flash steamed to power a turbine, which is connected to a generator, and produces electrical energy. The geothermal water is never exposed to the air, and is injected back into the periphery of the reservoir. The hot water from underground heats working fluid (e.g., a liquid organic compound) that has a lower boiling point than water (e.g., isobutane). The two liquids are kept completely separate through the use of a heat exchanger used to transfer the energy via heat from the geothermal water to the working fluid. The secondary fluid vaporizes into gaseous vapor and (like steam) the force of the expanding vapor turns the turbines that power the generators. In other words, the organic working liquid creates steam, which flows through a turbine and powers a generator to create electrical power. The water in the pipe is recycled back to the ground, to be re-heated by the Earth and provide future heat for the organic compound again. If the power system uses air cooling the geothermal fluids never make contact with the atmosphere before they are pumped back underground. These systems are the most recent development, and can accept fluid temperatures as low as 57 °C. The moderately hot geothermal water is passed by a secondary fluid with a much lower boiling point than water. This causes the secondary fluid to flash vaporize, which then drives the turbines. This is the most common type of geothermal electricity station being constructed today. Both Organic Rankine and Kalina cycles are used. The thermal efficiency of this type station is typically about 10–13%. In

binary systems, water is only used as a heating agent, and is not exposed or evaporated. It can be recycled, used for other purposes, or released into the atmosphere as non-toxic steam. In general, however, binary systems generally emit no visible steam or water vapor plumes. If the geothermal fluid is not contained and recycled in a pipe, it can absorb biologically toxic substances in the earth, such as arsenic, boron, selenium, mercury, and fluoride. These toxic substances can be carried to the surface and released when the water evaporates. In addition, if the fluid leaks to other underground water systems, it can contaminate clean sources of drinking water and aquatic habitats.

4. **Enhanced geothermal systems (EGS)** - EGS uses drilling, fracturing, and injection to provide fluid and permeability in areas that have hot, but dry, underground rock. To develop an EGS, an 'injection well' is drilled into the ground. Depending on the type of rock, this can be as shallow as 1 kilometer to as deep as 4.5 kilometers. High-pressure cold water is injected into the drilled space, which forces new fractures in the rock, expands existing fractures, and dissolves areas. This creates a reservoir for/of underground fluid. Water is pumped through the injection well, and absorbs the rocks' heat as it flows. This hot water, called 'brine', is then piped back up to earth's surface through a 'production well'. The heated brine is contained in a pipe. It warms a secondary [working] fluid that has a low boiling point, which evaporates to steam and powers a turbine. The brine cools off, and cycles back down through the injection well to absorb underground heat again. There are no gaseous emissions besides the water vapor from the evaporated liquid. Pumping water into the ground for EGSs can cause seismic activity.

   NOTE: *In 2009 an EGS Geothermal project in Basel, Switzerland, was cancelled after the injection process caused hundreds of tiny earthquakes that grew to more significant seismic activity even after the water injection was halted.*

5. **Hot water system (plant/station)** - This system uses a hot water reservoir that does not reach high enough temperatures to become steam, but is still a viable source of direct heating. The water is not used to produce electrical power, but is piped and exchanged to heat desired areas. The heat from the pipes (and venting systems) radiate heat into a surface and/or area. Pipes return the water to the hot water reservoir to be reheated and introduced back into the system. Applications include, but are not limited to: space heating and cooling; food preparation; hot spring bathing and spas (balneology); agriculture; aquaculture; greenhouses; snow melting; and production processes.

There are two rock-based geothermal power systems; one of which is designed specifically as a heating and cooling system, and the other is designed to produce electrical power:

1. **Geoexchange system (geothermal heat pump, GHP) for heating and cooling** - This is a type of heating and cooling system. Technically, this kind of system is not geothermal since it uses a combination of the ground's low relative heat and indirect solar energy, not the earth's geothermal energy. Geoexchange systems involve the drilling of a well 3 to 90 meters deep (shallower than most oil and natural gas wells). The system does not require the fracturing of bedrock. A pipe connected to a system is arranged in a continuous loop ("slinky loop") that generally circles underground and above ground. However, the loop can also be contained entirely underground, for instance, to heat a parking lot or landscaped area. In this system, a working fluid (such as water or glycerol, similar to a car's antifreeze) moves through the pipe (and accompanying heat exchanger and ductwork, if present). This system is mostly used for heating and cooling on a seasonal basis. During the cold season, the liquid absorbs underground thermal heat (from a geothermal and/or solar source). It carries the heat to its desired location. The pipe can be connected to the infrastructure of a building to give off heat into the building through a duct/heat exchange system. These heated pipes can also run through hot water tanks. During the warm season, the geoexchange system works the opposite way: the liquid in the pipes is warmed from the heat in the building (or ground heat source), and carries the heat to be cooled underground.

   NOTE: *Some animals burrow underground for warmth in the winter and to escape the heat of the summer.*

2. **Geoexchange system (geothermal heat pump, GHP) for electrical power** - Unlike geoexchange for heating and cooling, this system does use geothermal energy to do work, and must be placed in a location with a sufficiently high underground temperature. Geoexchange systems for electrical power are typically self-contained tubular units with two principal sections: the process section and the heat absorption section. The heat absorption

section is otherwise called the 'geothermal riser'. The riser has its own pump with an oil compound (biodegradable and non-toxic). The oil is sent downward through a coaxial system to where the geothermal heat exists. As the oil gets hot at the location of the thermal riser, it comes back up through the center tube. The hot oil is transferred to a heat exchange, which contains two chemicals (isopentane and isobutane). These chemicals are pressurized and in liquid form when they absorb thermal energy from the heat exchanger. Once they are heated from the exchange, they turn into a gas, which drives a turbine connected through a shaft to a generator, and produces electrical power. After the gas moves through the turbine it enters a condensing system, and becomes liquid again, which is then pumped back down to the heat exchanger. These systems typically require an operating temperature of at least 148˚C. The power generating system operates at ~10,000rpm and needs only 108˚C. Both sections of the device are closed cycles. One of the most well-known geoexchange electrical power systems is the Power Tube, which is available with the following specifications.

- 10 megawatt system = 55m in length, 142cm diameter (with a 30.48cm diameter riser).
- 5 megawatt system = 47m in length, 112cm diameter (with a 22.86cm diameter riser).
- 1 megawatt system = 43m in length, 92cm diameter (with a 12.7cm diameter riser).

There are also co-produced geothermal systems:

- **Co-produced geothermal power** - These systems use heat obtained from the steam and hot water produced as a by-product of petroleum and natural gas wells to power an electrical generator.

11.2.1 The Cooling subsystem

Most geothermal electrical power generating system include a cooling subsystem. Cooling can occur through air cooling (e.g., fans) and/or water/wet cooling (e.g., cooling towers and evaporative cooling). A cooling system is designed to prevent turbines from overheating and prolong system life.

> **NOTE:** *A 'cooling tower' is a heat rejection/ dissipation device that ejects waste heat to the atmosphere through the cooling of a water stream to a lower temperature.*

Wet/evaporative cooling used in water cooled systems requires a continuous supply of cooling water and creates vapour plumes (i.e., air emissions as water vapor emissions). Usually, some of the spent steam from the turbine (for flash- and steam-type plants) can be condensed for this purpose. Vapor plumes are unaesthetic and air cooled systems are preferred in areas where the viewshed is sensitive to the effects of vapor plumes. Vapor plumes can also introduce contaminants into the atmosphere, depending upon the purification of the water source.

Air cooled systems, in contrast to the relative stability of water cooled systems, can be extremely efficient in the winter months, but are less efficient in hotter seasons when the contrast between air and water temperature is reduced, so that air does not effectively cool the organic fluid. Air cooled systems emit no water vapor, and thus blend easily into the environment. Air cooled systems are beneficial in areas where extremely low emissions are desired, or in arid regions where water resources are limited, since no fluid needs to be evaporated for the cooling process.

11.2.2 Geothermal resource assessment

Different types of geothermal energy are available on different parts of the planet. Satellites are used to determine geothermal hotspots. Thermal satellite imagery usually provides a 10km x 10km view of the ground. With this information you can determine how large and/or how many geothermal systems can potentially be installed without extracting too much energy (i.e., without cooling the area down).

Geothermal plants are designed for a specific resource. In other words, after satellite imagery is analyzed, wells are drilled, and then, the thermal characteristics and material output (e.g., steam and/or water) of the well are measured. And then, a geothermal system is selected and designed for that specific resource.

The part of the planet known as "Iceland", for example, has abundant sources of hot, easily accessible underground water. Europe has a significant volume of hot dry rock as a source of geothermal energy.

11.2.3 Environmental impact

Before construction of a geothermal power generation system, an environmental assessment must be completed to determine potential social system and ecosystem effects.

11.2.3.1 Environmental contamination and disruption

- Drilling into the earth's crust produces 'sludge', which is often rich in zinc and sulfur (and other potential pollutants). If the sludge is directly released into the environment it can harm ecosystems.
- Water that flows through underground reservoirs can pick up trace amounts of toxic elements. These harmful substances can be leaked to water sources

or the atmosphere if the geothermal system is not properly insulated. Hence, the steam used at geothermal plants can become a source of air pollution if it is released into the atmosphere. Frequently it is heavily laced with salts and sulfur compounds that are leached from the earth's crust. If the steam is simply condensed and released into the natural waterways, the high levels of salts and sulfurs can be toxic to aquatic wildlife. If released into the air, the toxins, in the form of acid rain, can still find their way into surface water systems and kill aquatic animals. Geothermal plant operators often cool and condense the steam produced at their plants and recycle it into their wells for these reasons.

- Some geothermal reservoir fluids contain varying amounts of certain gases, including carbon dioxide, which may be emitted as steam into the environment (with any accompanying particulate matter). In concern to particulate matter, mercury for example, is not present in every geothermal resource. However, if mercury is present in a geothermal resource, using that resource for power production could result in mercury emissions, depending upon the technology used. Because binary plants pass geothermal fluid through a heat exchanger and then return all of it to the reservoir, binary plants do not emit any mercury.
- Natural geothermal fluids contain varying concentrations of potentially toxic minerals and other elements, and are extremely hot when they reach the surface of the Earth. For these reasons, geothermal fluids can be dangerous to humans and surrounding ecosystems. Hence, fluids from geothermal reservoirs are injected back into the earth and are not allowed to be released into surface waterways. However, geothermal effluents can sometimes (depending upon composition) be stored in evaporative ponds, rather than be injected back into the system. Again, depending upon composition, these ponds may be safe for bathing/swimming.
- During operation, there is the possibility of aquifer/groundwater contamination.
- Wells must often be dug, which may disrupt the natural flow of groundwater.
- Total non-condensable gas emissions (e.g., sulfur dioxide, methane, CO_2 etc.) from geothermal resources are calculated as a percentage of the total steam emitted (generally, less than 5%). Conversely, air emission from combustion (e.g., coal fired electrical plant) contain are a much higher percentage of emissions.

11.2.3.2 Subsidence

- Subsidence is the slow, downward sinking of a land surface. Other types of ground deformation include upward motion (inflation) and horizontal movements. In some cases, subsidence can damage infrastructure, such as roads, buildings and irrigation systems, or even cause tracts of land to become submerged by nearby bodies of water. Although it can occur naturally, subsidence can also occur as a result of the extraction of subsurface fluids, including groundwater, hydrocarbons, and geothermal fluids. In these cases, a reduction in reservoir pore pressure reduces the support for the reservoir rock itself and for the rock overlying the reservoir, potentially leading to a slow, downward deformation of the land surface. While subsidence can be induced by thermal contraction of the reservoir due to extraction and natural recharge, properly placed injection (see injection sections) reduces the potential for subsidence by maintaining reservoir pressures.
- Geothermal sights can experience subsidence (setting or sinking of land). Geothermal plants have been linked to subsidence, or the slow sinking of land. This happens as the underground fractures collapse upon themselves. In some areas of New Zealand, the ground under a geothermal power plant subsides at a rate of almost a half a meter every year. This can lead to damaged pipelines, roadways, buildings, and natural drainage systems.

11.2.3.3 Water Depletion

- Water-based geothermal system can/will deplete the naturally existing underground water (over time). In general, geothermal systems need additional replenishing sources of water, because the steam released exceeds the amount of water that naturally flows into the systems. To restore some of the former capacity, 'water injection' is used. Re-injecting water can sometimes help a cooling geothermal site last longer. However, this process can cause earthquakes. The process of injecting high-pressure streams of water into the Earth can result in minor seismic activity, or small earthquakes.
- In some geothermal power systems, when electrical power is generated [some amount of] steam is lost to evaporation.
- Depending upon how it is used and controlled, wastewater (from human or other animal waste) can be used as an alternative replenishing source of water. In other words, wastewater can be used for 'water injection'. Waste water injection projects serve the dual purpose of eliminating wastewater,

which would otherwise be dumped into local waterways, and rejuvenating geothermal reservoirs with new water sources.

11.2.3.4 Heat depletion

- Geothermal power is considered to be sustainable, because the heat extraction is small compared to the Earth's heat content. Extraction, however, must still be monitored to avoid local depletion. Although geothermal sites are capable of providing heat for many decades, individual wells may cool down or run out of water. Most wells that extract heat from the Earth will eventually cool, especially if heat is extracted more quickly than it is given time to replenish. Hence geothermal system require appropriate control/management over the amount of energy extracted.

11.2.3.5 Aesthetics

- Geothermal power plants can be designed to blend-in to their surrounding more so than other power generation methods, and they can be located on multiple-use land.

11.2.3.6 Travel

- The land around geothermal vents is often unstable, and unfit for human habitation. Hence, teams who work at geothermal stations may have to travel some distance from their place of habitation.

11.2.3.7 Seismicity

- Earthquake activity (seismicity) is generally caused by displacement across active faults in tectonically active zones. An earthquake occurs when a body of rock is ruptured and radiates seismic waves that shake the ground. Although it typically occurs naturally, seismicity has at times been induced by human activity, including the development of geothermal fields, through both production and injection operations.

12 Nuclear power

Nuclear power uses refined nuclear material, which has been turned into a fuel rod. The fuel rod is placed in water where the to heat created by fission turns the water into steam. The steam turns the blades of a steam turbine to produce electricity.

13 Energy from biomass and hydrocarbon

A.k.a., Biomass energy.

Hydrocarbons are different than carbohydrates. Hydrocarbons, per se, only contain hydrogen and carbon. The word "carbohydrate" means: carbo - carbon; hydr - hydrogen; ate - oxygen. Lipids (fats) are organic compounds that contain the same elements as carbohydrates: carbon, hydrogen, and oxygen. However, the hydrogen-to-oxygen ratio is always greater than 2:1. Proteins also contain carbon, hydrogen, and oxygen like carbohydrates and lipids, but they also contain nitrogen, and often, sulfur and phosphorus. All biomass is composed of some combination of carbohydrates, lipids, and/or proteins.

13.1 Biomass

Biomass is the term for organic matter derived from living organism, recently living organisms, and their waste that can be used as a source of energy. Biomass is "organic", meaning it is made of material that comes from living organisms and contains carbon atoms. Biomass may also be defined as biologically-produced matter based in carbon, hydrogen, and oxygen also. Biomass contains potential energy originating from the sun. Energy obtained from any biomass source is often called "bioenergy".

As an energy carrier ("source"), biomass can either be used directly via combustion to produce heat and light, or indirectly after converting (concentrating and purifying) it to various forms of biofuel. Biomass can be used to produce heat (thermal energy), light (as visible EM radiation), mechanical power, electrical power, and fuel (as well as compost). Biomass can be transformed into usable carriers of energy through direct and indirect means. Biomass can be burned/combusted to produce heat and light (direct), converted into electrical and mechanical power (directly through heat), or processed into biofuel (indirect). In other words, as an energy source, biomass can either be used directly to produce heat and light, or indirectly after converting it to various forms of biofuel (i.e., an energy storage medium).

> **NOTE:** *Biomass is thought to our oldest source of energy after the sun. Combustion (as fire) is understood to be the first controlled chemical reaction discovered by humans, in the form of campfires and bonfires. The ability to control fire (as a technology) led to dramatic changes in the habits (and potentially, the physiology) of early hominids. The combustion of biomass (plant material and animal waste) is the oldest known source of power production.*

Biomass generation is an integral part of Earth's carbon cycle. The carbon cycle is the process by which carbon is exchanged between all layers of the Earth: atmosphere, hydrosphere, biosphere, and lithosphere. Between periods of exchange, carbon is sequestered (stored as terrestrial/biologic or geologic). The carbon in fossil fuels has been sequestered for millions of years. Unlike plants, when hydrocarbons present in fossil fuels are extracted and burned for energy, their sequestered carbon is released into the atmosphere at a volume that cannot be immediately re-sequestered by the earth's natural carbon cycle. Fossil fuels take thousand years to form naturally. They cannot be replaced and sequestered as fast as they can be consumed.

In contrast to fossil fuels, biomass comes from recently living organisms. The carbon in biomass can continue to be exchanged by the carbon cycle. This is because biomass has a lower 'energy density' than fossil fuels. In order to ensure that the carbon cycle continues as trending and desired, biomass materials/resources must be sustainably farmed.

> **NOTE:** *Biomass can be combusted or converted to biofuel.*

13.2 Biomass sources

> **NOTE:** *Biomass is the same energy that makes fossil fuels. Fossil fuels are made through the concentration of biomass over time by heat and pressure within the earth.*

Technically, any plant or animal matter could be used as a source of biomass, because it will contain hydrogen, carbon, and oxygen. A wide variety of biomass forms are available on the planet, and biomass can be produced anywhere that plants or animals can live.

Well-known and highly effective sources of plant matter biomass include, but are not limited to: hemp, bamboo, corn, miscanthus, poplar, switchgrass, sorghum, sugarcane, and a variety of tree species, ranging from willow and eucalyptus, to oil palm (palm oil). Animal waste is another common form biomass.

Biomass resources can be classified as follows:

1. Biologically cultivated materials and residues (e.g., agricultural, aquaculture, permaculture, algae-culture etc.)
2. Forestry materials and residues
3. Food residues and by-products
4. Animal by-products (tallow, fish oil)
5. Animal/human solid waste
6. Landfill gases

13.2.1 Biofuel

Plant and animal matter, as an originating source of biomass, can be converted into biofuel. A **biofuel** is a fuel that is produced through natural biological processes, such as anaerobic digestion, rather than a fuel produced by geological processes, such as those involved in the formation of fossil fuels from prehistoric biological matter. Biofuels are classified into two primary categories:

1. First-generation biofuels - First-generation biofuels are derived from food-based biomass sources, which are fermented via sugars (carbohydrates/starch) present in the biomass to produce bioethanol. Bioethanol is an alcohol fuel that can be used directly in a fuel cell to produce electricity, or serve as an additive to gasoline.
2. Second-generation biofuels - Second-generation biofuels utilize non-food-based biomass sources. These biofuels mostly consist of lignocellulosic biomass, which is not edible. Lignocellulosic is a chemically inert and structural rigid form of biomass. Advanced second-generation biofuels are generally made from non-food feedstocks (feedstock that could not also be used as food), such as municipal waste, algae, perennial grasses, and wood chips. These fuels include cellulosic ethanol, bio-butanol, methane gas (biogas), methanol and a number of synthetic gasoline/diesel equivalents.

13.3 Biomass creation

NOTE: *The faster a plant grows, the more efficient it is as a source of biomass energy.*

Technically, all biomass (energy) is derived from plant- or algae-based material, which derives a significant portion of its energetic existence from solar electromagnetic radiation. Plants and other living organisms absorb the sun's energy, and use it to create the organic matter of their body. When biomass is burned, this stored energy is released as electromagnetic (and thermal) energy.

Through the use of solar electromagnetic radiation (energy), green plants convert carbon dioxide and water into carbohydrates and oxygen in a process called photosynthesis. It could be said that these plants "breath in" carbon dioxide and "breathing out" oxygen; however, in plants respiration is a different process. In green plants both photosynthesis and respiration occur as separate processes. During photosynthesis, water is broken down to form oxygen, and during cellular respiration, oxygen is combined with hydrogen to form water. Hence, plants, unlike other living things, produce oxygen. Essentially, photosynthesis is the opposite of [cellular] respiration.

DEFINITION: **Photosynthetically active radiation (PAR)** *designates the spectral range (wave band) of solar radiation from 400 to 700 nanometers that photosynthetic organisms are able to use in the process of photosynthesis. This spectral region corresponds more or less with the range of light visible to the human eye. Other living organisms, such as Cyanobacteria, purple bacteria and Heliobacteria, can exploit solar light in slightly extended spectral regions, such as the near-infrared. These bacteria live in environments such as the bottom of stagnant ponds, sediment, and ocean depths. Because of their pigments, they form colorful mats of green, red and purple.*

In relatively bright light, photosynthesis is the dominant process. During photosynthesis, chlorophyll, the pigment that makes plants green, absorbs electromagnetic radiation (energy) from the sun and uses it along with carbon dioxide and water to make carbohydrate molecules (sugars). Carbohydrates are complex compounds composed of carbon, hydrogen, and oxygen. At night, or in the absence of light, photosynthesis essentially ceases (in most plants), and respiration is the dominant process. However, plants respire continuously, day and night -- respiration does not depend on light. Whereas photosynthesis absorbs energy (from sunlight), aerobic respiration yields energy (as a result of the oxidation of glucose (a carbohydrate molecule, $C_6H_{12}0_6$).

1. Photosynthesis:
   - $6CO_2 + 6H_20$ + energy/energy in > $C_6H_{12}0_6 + 6O_2$
2. Aerobic respiration:
   - $C_6H_{12}0_6 + 6O_2 > 6CO_2 + 6H_20$ + energy/energy out

Biomass generation requires carbon dioxide; hence cultivated and combusted biomass in the appropriate ratio is carbon dioxide neutral. When the carbohydrates that plants have made are burned, they turn back into carbon dioxide and water, and release the energy they captured from the sun.

NOTE: *Some plants absorb carbon dioxide at during the day, and release oxygen at night. Typically, desert plants, such as cactus and other succulents, and epiphytic bromeliads and orchids from the jungle absorb CO_2 during the day and release oxygen at night. These plants include, but are not limited to snake plants, Easter and Christmas cactus, aloe vera plants (a species of succulent plant), mother-in-laws tongue, areca palm, a variety of bromeliads, and orchids such as the moth and the dendrobium.*

Plant-based biomass is often specifically called 'lignocellulosic biomass'. Lignocellulose is composed of carbohydrate polymers (cellulose, hemicellulose), and an aromatic polymer (lignin).

NOTE: *The efficiency of photosynthesis is low, about 5% maximum (solar energy to energy in sugar).*

In general, good biomass material is a combination of cellulose (~60%), lignin (~30%), and other organic materials (~10%). Cellulose and lignin start with simple sugars (glucose) made through photosynthesis. Cellulose is a fiberous organic compound used as structural material in plants (fibers). Cellulose is the most abundant organic polymer on Earth. Lignin is a class of complex organic polymers that form important structural materials in the support tissues of vascular plants and some algae.

Algae is a source of biomass energy. Some algaes produce energy through photosynthesis at a much higher rate than any other biomass source. Further, algae contains oils that can be converted into biofuels (and/or food). Algae can be fermented to produce biofuels such as ethanol, butanol, and methane, as well as biodiesel and hydrogen. Algae can be grown in salt and fresh water; it does not require soil. Algae takes up much less space to produce than other biomass/ biofuel crops; particularly because algae can be grown in bioreactors.

13.4 Biomass to biofuel conversion technologies

Biomass can be converted into a more concentrated energy carrier/source known as biofuel (also known as 'direct energy' or a 'concentrated energy resource'). Biofuels are easier to transport and are more energy dense than their original biomass resource. Also, their combustion characteristics are more convenient and predictable than raw biomass. The conversion of biomass to biofuel can be achieved by different methods, which are broadly classified into: thermo-chemical, biochemical, and chemical. The conversion process leads to biofuel that can take any of the following three forms of matter: biofuel solid; biofuel liquid; or biofuel gas.

1. **Thermo-chemical conversion** is the application of heat and chemical processes in the production of concentrated energy products from biomass (i.e., biofuel).
2. **Biochemical conversion** involves use of enzymes, bacteria, or other microorganisms to break down biomass into liquid fuels. The primary processes are anaerobic digestion/decomposition, fermentation, and composting.
3. **Chemical conversion** involves use of chemical agents to convert biomass into liquid fuels.

Note that hydrocarbons (as hydrocarbon fuels) can be derived from some biomass. There are a variety of ways of producing hydrocarbons from biomass since biomass is a mixture of carbon, hydrogen, and oxygen, and a pure hydrocarbon consists entirely of hydrogen and carbon. For instance, gasification converts whole biomass into a mixture of carbon monoxide and hydrogen gases, which can then be used to synthesize hydrocarbons.

13.4.1 Thermo-chemical conversion

Thermo-chemical conversion involves the use of heat as a significant mechanism for the chemical conversion of biomass into biofuel. Pyrolysis, torrefaction, and gasification are the basic thermochemical conversion technologies.

1. **Torrefaction -** Before biomass can be combusted, it must be dried. Torrefaction, like pyrolysis, is the conversion of biomass to a drier and refined form with the application of heat in the absence of oxygen, but at lower temperatures than those typically used in pyrolysis. During torrefaction, biomass is heated to about 200°C - 320°C. The biomass dries out so completely that it loses the ability to absorb moisture, or rot. It loses ~20% of its original mass, but retains 90% of its energy. The lost energy and mass can be used to fuel the torrefaction process. Torrefaction produces a solid biofuel. During torrefaction, biomass becomes a dry (generally, black) material. It is then compressed into briquettes. Biomass briquettes are hydrophobic, meaning they repel water. This makes it possible to store them in moist areas. The briquettes have a high energy density and are easy to burn.
2. **Pyrolysis** - Heating biomass to 200°C - 320°C in a pressurized environment without the presence (or with a very low presence) of oxygen. The absence of oxygen prevents combustion, causing a different chemical alteration to the substance. Pyrolysis produces a dark liquid called pyrolysis oil, a synthesis gas called syngas, and a solid residue called biochar.
    A. **Pyrolysis oil (bio-oil or biocrude)** - a type of tar (a mixture of hydrocarbons and free carbons). It can be combusted, used as a component in other fuels, and used as a compound in plastic.
    B. **Syngas (synthesis gas)** - a fuel gas mixture consisting primarily of hydrogen, carbon monoxide, and often, carbon dioxide. It can be converted into fuel, such as, synthetic natural gas. And, it can be converted into methane (used as a replacement for natural gas).
    C. **Biochar** - a type of charcoal, which consists of carbon and any remaining ash, obtained by removing water and other volatile constituents from animal and vegetation substances. Biochar is high in carbon. It is used as a soil amendment/conditioner. Biochar enriches soil and prevents it from leaching pesticides and other nutrients into runoff. Biochar is also an excellent carbon sink. Carbon sinks are reservoirs for carbon-containing chemicals, including greenhouse gases.
    D. **Gasification** - A process that converts organic or fossil fuel based carbonaceous materials into carbon monoxide, hydrogen, and carbon

dioxide, producing syngas and slag. During the gasification process, a biomass feedstock is heated to more than 700° C with a controlled amount of oxygen. During gasification, syngas is purified of sulfur, mercury, and other polluting particulates. Slag, a by-product of the process, forms as a glassy, molten liquid, which can be used to make shingles, cement, or asphalt.

13.4.2 Biochemical conversion

Many highly efficient biochemical processes have developed in nature to break down the molecules of which biomass is composed. Biochemical conversion processes use the enzymes of bacteria and other microorganisms (e.g., fungi) to convert biomass to gas and liquid fuels. The primary processes are anaerobic digestion/decomposition, fermentation, and composting:

1. **Anaerobic decomposition/digestion to methane** - the process by which microorganisms break down organic material in an oxygen-free (or low oxygen) environment. Anaerobic digestion is widely used for the production of methane- and carbon-rich biogas from biomass (e.g., crop residues, food scraps, and human and animal waste). Anaerobic digestion is frequently used in the treatment of wastewater, and to reduce emissions from landfills. Controlled anaerobic digestion is usually a multi-stage process. First, the carbohydrates are broken down. The resulting sugars and amino acids are then converted into carbon dioxide, hydrogen, ammonia, and organic acids. Finally, these products are converted into methane and carbon dioxide. Microorganisms are used in each stage of the process. These mixed cultures allow digesters to be operated over a wide temperature range, for example, above 0° C and up to 60° C. Solid remnants of the original biomass input are left over after the digestion process. This by-product, or digestate, has many potential uses. Potential uses include fertilizer (although it should be chemically assessed for toxicity and growth-inhibiting factors first), animal bedding and low-grade building products like fiberboard. Methane production is most useful with animal and human wastes, as well as landfill wastes, where it happens naturally.
2. **Fermentation to biofuel** - a biological/metabolic process by microorganisms that converts carbohydrate (sugar) to acids, gases, or alcohol (e.g., ethanol). Fermentation generally involves multiple stages. The science of fermentation is known as zymology. Biomass is the only short-term renewable energy source that can be converted into liquid biofuels, such as ethanol and biodiesel. Ethanol is made by fermenting biomass that is high in carbohydrates, such as hemp, sugar cane, wheat, or corn through ethanol/alcoholic fermentation. Biodiesel is made from combining ethanol with animal fat, recycled cooking fat, or vegetable oil. Biofuels do not combust as efficiently as petrol. They can, however, be blended with petrol.
   A. **Ethanol fermentation** (alcoholic fermentation) - converts sugars such as glucose, fructose, and sucrose into cellular energy, producing ethanol and carbon dioxide as a by-product. Note that as a material resource, ethanol can be used as a consumable, as a disinfectant, and as a fuel.
3. **Composting** - the process by which organic matter is decomposed and recycled as a fertilizer and soil amendment. Bacteria requiring oxygen to function (aerobic bacteria) and fungi manage the chemical process by converting the inputs into heat, carbon dioxide and ammonium. Note that the composting process naturally produces heat (thermal energy).

13.4.3 Chemical conversion

The utilization of chemical processes to convert biomass into biofuel.

1. **Transesterification** - Transesterification is a chemical reaction through which fatty acids from oils, fats, and greases are bonded to alcohol. This process reduces the viscosity of the fatty acids and makes them combustible. It is an organic chemistry process wherein the organic group R" of an ester is exchanged with the organic group R' of an alcohol. The process refers to a reaction between an ester of one alcohol and a second alcohol to form an ester of the second alcohol and an alcohol from the original ester. It is the most common form of chemical-based conversion. Biodiesel is a common end-product of transesterification, as are glycerin and soaps. For instance, mixing methanol with sodium hydroxide will create sodium methoxide. This liquid can then be mixed into vegetable oil. When the mixture settles, glycerin is left on the bottom and methyl esters, or biodiesel, is left on top. The glycerin can be used to make soap (or many other products), and the methyl esters are washed and filtered. Almost any bio-oil (e.g., soybean oil), animal fat or tallow, or tree oil can be converted to biodiesel. Transesterification of algal oil is frequently accomplished with ethanol and sodium ethanolate serving as the catalyst.
2. **Black liquor production** - The kraft process (also known as kraft pulping or sulfate process) is a mechanical-chemical process that converts wood into the main component of paper using a hot

mixture of water, sodium hydroxide, and sodium sulfide, known as white liquor, that breaks the bonds that link lignin, hemicellulose, and cellulose. One of the by-products of the process is known as “black liquor”, which retains more than 50% of the wood’s biomass energy. Black liquor can be used as a biofuel feedstock.

3. **Hydrogen production** - Biomass has a high concentration of hydrogen, which can be chemically extracted and used as an energy source (e.g., hydrogen fuel cells).

13.5 Hydrocarbons

> **CLARIFICATION:** *Hydrocarbons are always named based on the longest carbon chain. When a hydrocarbon has a double bond we replace the -ane ending with -ene. When the hydrocarbon has more than three carbon the position of the double bond must be specified with a number. Hydrocarbons with triple bonds are named basically the same, we replace the -ane ending with -yne. Cyclic hydrocarbons with delocalized bonds are called aromatic hydrocarbons the most common of these is benzene.*

Hydrocarbon molecules have a high energy density relative to other molecules, and are “easy” to store and transport. Hydrocarbons are the simplest form of organic [lipid] compound, and contain only carbon and hydrogen (i.e., hydrogen + carbon only).

1. Biomass hydrocarbons
2. Fossil hydrocarbons

Hydrocarbons can be found in the forms of matter:

1. Gases (e.g. methane and propane)
2. Liquids (e.g. hexane and benzene)
3. Waxes or low melting solids (e.g. paraffin wax and naphthalene)
4. Polymers (e.g. polyethylene, polypropylene and polystyrene)

Hydrocarbons can be classified at a top-level as either open chain or closed chain:

1. **Aliphatic hydrocarbons (open chain)** - - formed by successive bonds between carbon atoms and may be branched or unbranched (linear or “normal”). Herein, the overall geometry of the molecule is altered by the different geometries of single, double, and triple covalent bonds. A normal/linear/unbranched hydrocarbon has one chain of consecutively bonded carbon atoms. A branched hydrocarbon has at least one carbon atom not bonded to the end carbon of a chain of consecutively bonded carbon atoms. Instead, at least one carbon atom forms a bond to an inner carbon atom in the chain of consecutively bonded carbon atoms. Aliphatic hydrocarbons (open chain) can be either saturated or unsaturated hydrocarbons, neither of which contain a benzene ring. Aliphatic hydrocarbons are classified based on the structure and bonding of the carbon skeleton into one of three groups: alkanes (saturated); alkenes (unsaturated double bond), and alkynes (unsaturated triple bond). Saturated hydrocarbons consist entirely of single bonds, wherein each carbon atom is connected to four other atoms. Unsaturated hydrocarbons have one or more double or triple bonds between carbon atoms. Aliphatic hydrocarbons tend to be flammable; they combust (undergo transformation) through which old bonds are broken, and new bonds are formed.

A. **Alkanes (a.k.a., saturated hydrocarbons; paraffins)** - Alkane molecules are those chemical structures that are based on carbon atoms having only single bonds, and that are completely saturated with hydrogen atoms. Each carbon atom is connected to four other atoms; either another carbon within the skeletal structure, or a hydrogen atom. Saturated hydrocarbons are the basis of petroleum fuels and are found as either linear or branched species. Alkanes can be described by the formula: C_nH_{2n+2}, where ‘n’ is the number of carbon atoms present.

B. Methane, where n=1, described by: CH_4

C. **Alkenes (a.k.a., olefins; unsaturated form of hydrocarbon)** - Alkene molecules contain at least one carbon-carbon double bond. Alkenes can be described by the one double bond formula: C_nH_{2n} (assuming non-cyclic structures). Alkene hydrocarbons are present in most organic and biological molecules. Alkene compounds do not occur naturally in crude oil, but are produced by reaction during the refining process. Example alkenes include: ethylene; propylene; or butylene.

D. **β-Carotene carotenoids** (a.k.a., tetraterpenoids and terpenoids) - Carotenoids are organic pigments that are found in the chloroplasts and chromoplasts of plants and some other photosynthetic organisms, including some bacteria and some fungi. Carotenoids can be produced from fats and other basic organic metabolic building blocks by all these organisms. There are over 600 known carotenoids, which are classified into two classes: xanthophylls (which contain oxygen) and carotenes (which are purely hydrocarbons, and contain no oxygen). Carotenes typically contain only carbon and hydrogen (i.e., are hydrocarbons), and are in the subclass of unsaturated hydrocarbons. The term, carotenoid, is a misnomer and originates from a scientist (1831) who proposed the term "carotene" for the hydrocarbon pigment he had crystallized from carrot roots.

E. **Alkynes (unsaturated form of hydrocarbon)** - Alkyne molecules contain at least one carbon-carbon triple bond. Alkynes can be described by the one triple bond formula: C_nH_{2n-2}. Alkyne hydrocarbons rarely occur in biological molecules or pathways.

2. **Cyclic hydrocarbons (closed chain)** - formed by successive rings of carbon. Whenever the ends of a carbon chain are joined together, that molecule is said to be cyclic.
   A. **Cycloalkanes** (cycloparaffins or naphthenes, distinct from naphthalene) - are the cyclic analog of an alkane. Cycloalkanes are alkanes that consist entirely of single bonds with at least three carbon atoms linked together to form a structural ring (hence, the prefix 'cyclo-'). In other words, cycloalkanes are alkanes in which all or some of the carbon atoms are arranged in a ring. Cycloalkanes are monocyclic saturated hydrocarbons, and hence, are arranged in a structure containing a single ring (monocyclic, possibly with side chains), and all of the carbon-carbon bonds are single. Note here that there are also polycyclic alkanes, which are molecules that contain two or more monocyclic cycloalkanes that are joined, forming multiple rings. If the carbon chain that forms the backbone of a straight-chain hydrocarbon is long enough, we can envision the two ends coming together to form a cycloalkane. One hydrogen atom has to be removed from each end of the hydrocarbon chain to form the CC bond that closes the ring. Cycloalkanes therefore have two less hydrogen atoms than the parent alkane. When a cycloalkane contains only one ring, the general formula is C_nH_{2n}. However, the complete chemical formula for cycloalkanes is $C_nH_{2(n+1-r)}$, where n is the number of carbon atoms and r is the number of rings. Cycloalkanes are named analogously to their normal alkane parent counterpart of the same carbon count: cyclobutane; cyclopropan; cyclobutane; cyclopentane; cyclohexane; etc. The larger cycloalkanes, with more than 20 carbon atoms are typically called cycloparaffins.
   B. **Cycloalkenes** - are the cyclic analog of an alkene. Cycloalkenes are alkenes that consist of three or more carbon atoms linked together with at least one carbon-carbon double bond to form a structural ring (hence the prefix 'cyclo-'). They have no aromatic character.
   C. **Cycloalkynes** - are the cyclic analog of an alkyne. A cycloalkyne consists of a closed ring of carbon atoms containing one or more triple bonds. Cycloalkynes have a general formula C_nH_{2n-4}.
   D. **Aromatic hydrocarbons (arenes or aryl hydrocarbon)** - have at least one benzene-like ring (i.e., aromatic ring) of alternate single and double bonds with delocalized pi electrons between carbon atoms forming the ring(s). A benzene ring is a ring of six carbons with alternating double and single bonds. As a result, the benzene has six hydrogens and the formula for a benzene molecule is C_6H_6. Aromatic hydrocarbons can be monocyclic (MAH) or polycyclic (PAH). These compounds possess unique properties due to the delocalized electron density in benzene, including additional stabilization. Note that the term 'aromatic', and was assigned before the physical mechanism determining aromaticity was discovered; the term was coined because many of the compounds have a sweet or pleasant odour.

Some sources of biomass contain hydrocarbon substances known as terpenoids, which are similar to petroleum. These plants and algaes are known as "hydrocarbon plants" (or "petro-plants") and "hydrocarbon algaes". Hydrocarbon plants use unique metabolic pathways to produce hydrocarbon products. For instance, some of these plants produce a type of natural rubber (e.g., latex) that contains liquid hydrocarbon terpenoids of a high molecular weight, which can be converted into fuel and other products. Natural rubber is a hydrocarbon that contains long chains of alternating C=C double bonds and C-C single bonds. Terpenoid hydrocarbons may be extracted from the bulk matter of such plants through the use of organic solvents.

Well known families of hydrocarbon plants include, but are not limited to: Apocynaceae, Asclepiadaceae, Dipterocarpaceae, Euphorbiaceae, Hardwickia Pinnata (family Leguminosae), Moraceae, Sapotaceae, and sunflower (family Composiae). The latex of Euphorbia Lathyrus contains a fairly high percentage of terpenoids. And, the carbohydrate (hexose) from such plants can be used for ethanol formation.

13.6 Power from biomass, fossil fuels, and other hydrocarbons

Combustion is the primary way by which power is produced from biomass, biofuel, fossil fuel, and other hydrocarbons. Biomass (and its refined products) can be ignited and combusted for heat, light, and fluid [gas] pressure. Take note that when gases are combusted they generally expand irreversibly, and the fluid pressure which has been generated may be used to power a turbine or an engine.

1. **Combustion as direct firing/burning** - the direct burning of biomass or biofuel in the presence of oxygen. Fire can be used for heat, light, and/or fluid [gas] pressure.
   A. In a furnace, biomass burns in a combustion chamber converting the biomass into heat. The heat may be distributed in the form of hot air or water. A common type of furnace for area heating is known as a wood-pellet stove. A pellet stove is a stove that burns compressed wood or biomass pellets to create a source of heat. Wood pellets are the most common type of pellet fuel and are generally made from compacted sawdust and related industrial wastes from the milling of lumber, manufacture of wood products and furniture, and construction. The biomass (possibly in the form of pellets) are placed in a hopper, which feeds the mass into a furnace, where it is burned. The heat may used to boil water in a boiler. In a boiler, the heat of combustion is converted into steam. Steam can be used to produce mechanical energy through a turbine, electrical energy through a turbine-generator, or heating.
   B. In an engine, combustion occurs as a flame that propagates in a cylinder.
- **Combustion as co-firing/co-generation** - Biomass is combusted with a fossil fuel, often in pre-existing fossil fuel (coal) plants. Biomass can also be used in co-generation (a.k.a., combined heat and power, CHP), which is the simultaneous production of heat and electricity.

Some biofuel technologies can be directly converted into electrical power via electrochemical oxidation of the material. This electrochemical process can occur in carbon fuel cells, ethanol fuel cells, and microbial fuel cells. The fuel can also be consumed indirectly via a fuel cell system containing a reformer which converts the bio-mass into a mixture of CO and H_2 before it is consumed in the fuel cell.

14 Energy storage (secondary energy carriers)

> **CLARIFICATION:** *Primary energy carriers include all of the natural resources like natural gas, crude oil, coal, uranium, solar radiation, wind power, hydropower, and geothermal energy. Secondary energy carriers are those carriers of for which the production of other energy was needed (i.e., they required charging/producing by another power source).*

Energy can be stored in a variety of different carriers. There are many different types of energy stored in materials, and it takes a particular type of reaction to release each type of energy. In order of the typical magnitude of the energy released, these types of reactions are: nuclear, chemical, electrochemical, and electrical.

> **NOTE:** *Stored energy can be converted to power through the use of appropriate conversion/ release technology.*

Energy storage refers to the storage of useful energy that may either be used directly or transmitted is input in an end-use application.

Energy can be stored only as potential energy. Storing energy requires mass. Today, the most common way of storing energy is in [a mass of] batteries. Energy storage usually means batteries, but there are other ways, like pumped hydro and molten salt. But whatever the technology, there are two primary performance parameters in concern to energy storage. And, the usefulness of an energy storage system depends on both of these quantities. The two quantitative parameters are:

1. How much total energy can the system store? (Think watt-hours)
2. How much power can it deliver at any moment? (Think watts)

Storage systems have to be able to store enough energy to last through the "blackout" periods, and they have to be able to deliver that energy fast enough to meet the electrical load.

Once you know both the energy storage capacity (say, in megawatt-hours) and the output power (say, megawatts), you can simply divide these numbers to find how long the backup power will last. For example, a 20 megawatt-hour storage facility delivering power at the rate of 2 megawatts will last for 20 ÷ 2, or 10 hours on a full charge.

- Fuel - Matter that stores energy is called a '**fuel**'. Materials that store energy for work are called fuels. The amount of energy a fuel or other energy carrying source contains is called 'energy density'. To acquire energy, you must use energy.

The system uses all available energy from locally generated sources (such as photovoltaic cells) first, then 'fills in' with power from the grid or, when the grid is not available, from batteries.

> **NOTE:** *The external energy of a collection of matter, or system, is related to the relative condition of the matter with respect to its environment.*

14.1 Measurement for energy storage

In general, energy storage is measured in Joules. The formula for energy storage is as follows:

- $J = \frac{1}{2}CE^2$ = 1 watt/second
- where, J = joules, C = farads and E = voltage of the charge.

14.2 Energy storage performance parameters

Energy is a quantity that can be calculated for a given static carrier. Therein, **capacity** is the measure of a system's potential to generate power (or in the case of batteries, both generate power and store energy).

All energy storage devices have:

1. A power rating
2. An energy capacity rating
   There are two main things to consider with the choice of energy storage. Directly,
3. Can it produce enough current to the application (e.g., a motor)?
4. Does it have enough stored energy to last a required amount of time (e.g., 1 hour)?

Batteries have:

- Capacity = amp-hours

Energy storage usually means batteries, but there are other ways, like pumped hydro and molten salt. But whatever the technology, there are two performance parameters of interest:

1. How much total energy can the system store? (Think watt-hours)
2. How much power can it deliver at any moment? (Think watts)
3. How much recover efficiency can the system restore

Once you know both the energy storage capacity (say, in megawatt-hours) and the output power (say, megawatts), you can simply divide these numbers to find how long the backup power will last.

There are two performance parameters of interest in energy storage:

1. How much total energy can the system store? (Think watt-hours)
2. How much power can it deliver at any moment? (Think watts)

The usefulness of a storage system depends on both of these quantities. A system that stored an enormous amount of energy wouldn't be very useful if it could only return that energy a few watts at a time. And a system powerful enough to light up a whole city wouldn't be good for much if its batteries died after a few minutes.

14.3 Carriers/sources/modes of energy storage (energy storage systems)

A widely-used approach for classifying EES systems is the determination according to the source/carrier/form of energy used. EES systems are classified into mechanical, electrochemical, chemical, electrical and thermal energy storage systems:
Take note that these energy storage systems are sometimes as [secondary] energy carriers.

14.3.1 Mechanical storage systems

Energy may be "stored" in a mechanical system. There are many different types of mechanical systems utilized for the storage of [mechanical] energy.

1. **Pumped hydro storage** - the storage of water at elevation, which is released when power is required.
2. **Compressed/pressurized air storage** - the storage of compressed gas, which is released when power is required.
3. **Flywheel energy storage** - the storage of rotational mechanical/kinetic energy in an accelerated rotor, which is released when power is required.
4. **Gravitational potential energy** - the storage of mass at elevation, which is released when power is required.
5. **Tension** - the storage of energy in a device that holds tension, which is released when power is required.

14.3.2 Pumped hydro storage (PHS)

> **NOTE:** *Technically, PHS and hydroelectric dam systems store gravitational potential energy - fluids stored at elevation.*

Pumped hydro storage is a method of keeping water in reserve at elevation. The water is pumped to a storage pool above the power generation mechanism at a time when power demand is low, such as during the middle of the night. The water is then allowed to flow back through the turbine-generators at times when demand is high and a heavy load is place on the system.

The reservoir acts much like a battery, storing power in the form of water at elevation when demands are low, and producing maximum power during daily and seasonal peak periods. An advantage of pumped storage is that hydroelectric generating units are able to start up quickly and make rapid adjustments in output.

Conventional pumped hydro storage systems use two water reservoirs at different elevations to pump water during off-peak hours from the lower to the upper reservoir (charging). When required, the water flows back from the upper to the lower reservoir, powering a turbine with a generator to produce electricity (discharging). There are different options for the upper and lower reservoirs. For example, higher elevation dams and ponds can be used as pumped hydro storage plants. For the lower reservoir, flooded mine shafts, underground cavities, and even the open sea are also technically possible. PHS has existed for a long time – the first pumped hydro storage plants were used in Italy and Switzerland in the 1890s. Advantages are the very long lifetime and practically unlimited cycle stability of the installation. Main drawbacks are the dependence on topographical conditions and large land use. The main applications are for energy management via time shift, namely non-spinning reserve and supply reserve.

14.3.3 Compressed air (compressed gas) energy storage (CAES), also pressurized air storage

This is an electro-mechanical storage solution where air (or gas) is compressed with electrical power, and can then be released again to drive a power generator. Electricity is used to compress air (or other gas) and store it in either an underground structure or an above-ground system of vessels or pipes. When needed the compressed air may be used for fluid power, or it may be mixed with natural gas, burned and expanded in a modified gas turbine to produce mechanical and electrical power. Typical underground storage options are caverns, aquifers or abandoned mines. If the heat released during compression is dissipated by cooling and not stored, the air must be reheated prior to expansion in the turbine. This process is called diabatic CAES and results in low round-trip efficiencies of less than 50%. Diabatic technology is well-proven; the plants have a high reliability and are capable of starting without extraneous power. The advantage of CAES is its large capacity; disadvantages are low round-trip efficiency and geographic limitation of locations.

Compression of air creates heat; the air is warmer after compression. Expansion requires heat. If no extra heat is added, the air will be much colder after expansion. If the heat generated during compression can be stored and used during expansion, the efficiency of the storage improves considerably. There are three ways in which a CAES system can deal with the heat. Air storage can be adiabatic, diabatic, or isothermal.

1. **Adiabatic compressed air energy storage (ACAES)** - Adiabatic storage continues to keep the heat produced by compression and returns it to the air when the air is expanded to generate power.
2. **Diabatic compressed air energy storage (DCAES)** - Diabatic storage dissipates much of the heat of compression with intercoolers (thus approaching isothermal compression) into the atmosphere as waste; essentially wasting, thereby, the renewable energy used to perform the work of compression. Upon removal from storage, the temperature of this compressed air is the one indicator of the amount of stored energy that remains in this air. Consequently, if the air temperature is low for the energy recovery process, the air must be substantially re-heated prior to expansion in the turbine to power a generator.
3. **Isothermal compressed air energy storage (ICAES)** - Isothermal compression and expansion approaches attempt to maintain operating temperature by constant heat exchange to the environment.

14.3.4 Flywheel energy storage (FES)

In flywheel energy storage, rotational energy is stored in an accelerated rotor, a massive rotating cylinder.

This type of energy storage system is also sometimes known as a mechanical battery. The main components of a flywheel are the rotating body/cylinder (comprised of a rim attached to a shaft) in a compartment, the bearings and the transmission device (motor/generator mounted onto the stator). The energy is maintained in the flywheel by keeping the rotating body at a constant speed. An increase in the speed results in a higher amount of energy stored. To accelerate the flywheel electricity is supplied by a transmission device. If the flywheel's rotational speed is reduced electricity may be extracted from the system by the same transmission device. Advanced FES systems have rotors made of high-strength carbon filaments, suspended by magnetic bearings, and spinning at speeds from 20000 to over 50000 rpm in a vacuum enclosure. The main features of flywheels are the excellent cycle stability and a long life, little maintenance, high power density and the use of environmentally inert material. However, flywheels have a high level of self-discharge due to air resistance and bearing losses and suffer from low current efficiency.

14.3.5 Gravitational potential energy storage with solid mass

Changing the altitude of solid masses can store via an elevating system driven by an electric motor/generator. When power (energy) is required, the mass is released and pulled ("falls") toward the center of gravity.

14.3.6 Spring-Tension energy storage

Energy can be stored in a tension device, such as the winding of a spring in an pocket watch. Potential energy is stored in the spring's tension. When the tension on the spring is released, mechanical/kinetic energy (power) is released.

14.3.7 Chemical energy storage systems (secondary energy carriers)

The energy is stored in a static chemical medium (i.e., chemical fuel), such as petroleum fuel (e.g., coal, gasoline, diesel fuel, natural gas) or biofuel (e.g., biogas, char).

14.3.8 Solid fuel energy storage

Any solid fuel is a chemical energy storage medium.

14.3.9 Liquid fuel energy storage (a.k.a., power to liquid)

Any liquid fuel is a chemical energy storage medium.

14.3.10 Gaseous fuel energy storage (a.k.a., power to gas)

Power to gas refers to technology that converts electricity into a gaseous fuel such as hydrogen or methane. The main purpose of the following chemical energy storage systems is to use "excess" electricity (i.e., electricity when available) to produce hydrogen via water electrolysis. Once hydrogen is produced different ways are available for using it as an energy source/carrier, either as pure hydrogen or as SNG.

These chemical energy storage systems allow for the storage of large amounts of energy, up to the TWh range, and for greater periods of time – even as seasonal storage. Another advantage of hydrogen and SNG is that these universal energy carriers can be used in different habitat service sub-systems.

14.3.10.1 Hydrogen (H_2)

Hydrogen is the lightest of all chemical elements. It is odourless, colourless, and non-toxic. It has a high diffusibility, and boiling point of approximately -259.2°C, with ignition limits in air 4.0-75.0 Vol.%. Hydrogen can be stored as either compressed hydrogen (CH_2) in tanks at 200 bar and up to 700 bar in the near future, or as liquid hydrogen in cryogenic tanks, or as metal hydride. In order to increase storage density, hydrogen can be liquified. It is then called LH_2 (Liquid/Liquefied Hydrogen) and has to be stored and transported at -253°C (-423°F) in cryogenic tanks. Liquefaction is highly energy-consuming, and requires about one third of the energy content of the liquid hydrogen. Different approaches exist to storing the hydrogen, either as a gas under high pressure, a liquid at very low temperature, adsorbed on metal hydrides or chemically bonded in complex hydrides. However, for stationary applications gaseous storage under high pressure is the most popular choice. Smaller amounts of hydrogen can be stored in above-ground tanks or bottles under pressures up to 900 bar. For larger amounts of hydrogen, underground piping systems or even salt caverns with several 100 000 m^3 volumes under pressures up to 200 bar can be used.

Hydrogen can be produced via water-electrolysis. Electrolysis itself is the separation of bonded chemical elements. Therein, water-electrolysis involves running an electric current through water to split the bonded chemical elements into its compounds hydrogen and oxygen, both of which may then be stored.

Hydrogen may be used in fuel cells for local electricity generation. Therein, a typical hydrogen storage system consists of an electrolyzer, a hydrogen storage tank, and a fuel cell. An electrolyzer is an electrochemical converter which splits water with the help of electricity into hydrogen and oxygen. It is an endothermal process (i.e. heat is required during the reaction). Hydrogen is stored under pressure in gas bottles or tanks, and this can be done practically for an unlimited time. To generate electricity, both gases flow into the fuel cell where an electrochemical reaction, which is the reverse of water splitting (i.e., reverse of electrolysis) takes place: hydrogen and oxygen react and produce water, heat is released, and electricity is generated. For hydrogen storage systems specifically, the oxygen is generally vented to the atmosphere on electrolysis, and oxygen is taken from the air for the power generation.

> **NOTE:** *In fuel cells electricity is generated by oxidizing hydrogen or methane. This combined electrolysis-fuel cell process is an electrochemical energy storage system. However, both gases are multi-purpose energy carriers. Electricity can be generated in a gas or steam turbine.*

14.3.10.2 Oxygen (O_2)

Oxygen can be produced via water-electrolysis. Electrolysis itself is the separation of bonded chemical elements. Therein, water-electrolysis involves running an electric current through water to split the bonded chemical elements into its compounds hydrogen and oxygen, both of which may then be stored.

The oxygen can be compressed and stored in a storage tank. It can then be combusted to produce heat or its combustion can be used to power a turbine generator to produce mechanical and electrical power.

14.3.10.3 Methane and synthetic natural gas (SNG)

Methane is the simplest hydrocarbon with the molecular formula CH_4. Methane is more easily stored than hydrogen and the transportation.

Synthesis of methane (also called synthetic natural gas, SNG or syngas) is the second option to store electricity as chemical energy. Synthetic natural gas (syngas or SNG) can be created in a multi-step process, starting with hydrogen and oxygen. Hydrogen is then reacted with carbon dioxide in a Sabatier process, producing methane and water. Here a second step is required beyond the water splitting process in an electrolyzer, a step in which hydrogen and carbon dioxide react to methane in a methanation reactor. As is the case for hydrogen, the SNG produced can be stored in pressure tanks, underground, or fed directly into a gas grid. Several CO_2 sources are conceivable for the methanation process, such as fossil-fuelled power stations, manufacturing/ production installations, or biogas plants.

14.3.11 Biological energy storage

Technically, energy can be stored in biological systems via glycogen, starch, and lipid production. Mammals, for example, store energy in the form of fat (i.e., lipids) and liver glycogen. Many plants store energy in the form of starch.

14.3.12 Electrochemical storage systems

Energy may be stored within an electrochemical system.

> **NOTE:** *An* **uninterruptible power supply (UPS)** *is a device that allows an electrical (or other) system to keep running for at least a short time when the primary power source is lost. It also provides protection from power surges.*

Technically, a battery is a chemical potential energy storage system, however, they are classified herein under electrochemical storage because when the energy in a battery is released it is released into an electrical system. In thermodynamics, 'chemical potential' is defined as the time rate of change of internal energy of a system through changes in the number of particles in the system (or in the limiting case, the derivative of internal energy through the number of particles).

Chemical potential is characterized by the following abilities to do work in a chemical system:

1. To react with other substances (chemical reaction).
2. To move to another state (phase transition).
3. To reallocate the space (diffusion).

There are two types of non-flow batteries: primary batteries (disposable batteries), which are designed to be used once and discarded, and secondary batteries (rechargeable batteries), which are designed to be recharged and used multiple times.

Batteries come in many sizes, voltages, and amperes.

14.3.13 Primary batteries (non-rechargeable)

A primary battery is a portable voltaic cell that is not rechargeable. These batteries must be re-cycled after a single use. In general, the electrochemical reaction occurring in the cell is not reversible, rendering the cell non-rechargeable. As a primary cell is used, chemical reactions in the battery use up the chemicals that generate the power; when they are gone, the battery stops producing electricity and is useless. Primary batteries are used when charging is impractical or impossible (e.g., pacemaker). Primary batteries are designed for high specific energy, long storage times, instant readiness and/or long usage times.

The most common types of primary battery are listed below, but note that secondary batteries, such as lithium ion batteries, can be designed to function as a primary, and not secondary, battery.

14.3.13.1 Zinc Manganese oxide (alkaline battery)

Alkaline batteries are dependent upon a chemical reaction between zinc and manganese (IV) oxide (Zn/ MnO_2). The alkaline battery gets its name because it has the alkaline electrolyte, potassium hydroxide, instead of the acidic ammonium chloride or zinc chloride electrolyte of the zinc-carbon batteries. Other battery systems also use alkaline electrolytes, but they use different active materials for the electrodes.

14.3.13.2 Zinc-carbon (Leclanche)

A zinc–carbon battery is a dry cell battery that delivers a potential of 1.5 volts between a zinc metal electrode and a carbon rod from an electrochemical reaction between zinc and manganese dioxide mediated by a suitable electrolyte.

14.3.14 Secondary batteries (rechargeable)

A rechargeable battery, storage battery, secondary cell, or accumulator is a type of electrical battery that can be charged, discharged into a load, and recharged many times, while a non-rechargeable or primary battery is supplied fully charged, and discarded once discharged. In a secondary cell, the reaction can be reversed by running a current into the cell with a battery charger to recharge it, regenerating the chemical reactants.

In conventional secondary batteries, the energy is charged and discharged in the active masses of the electrodes.

The most common types of secondary battery are listed as follows:

14.3.14.1 Lead acid (LA)

Lead acid batteries are the world's most widely used battery type. Typical service life is 6 to 15 years with a cycle life of 1500 cycles at 80% depth of discharge, and they achieve cycle efficiency levels of around 80% to 90%. There are many sub-types of lead acid batteries. One disadvantage of lead acid batteries is usable capacity decrease when high power is discharged. For example, if a battery is discharged in one hour, only about 50% to 70% of the rated capacity is available. Other drawbacks are lower energy density and the use of lead, a hazardous material. LA systems are easy to recycle and the charging technology is simple.

14.3.14.2 Nickel cadmium and nickel metal hydride (NiCd, NiMh)

Compared to lead acid batteries, nickel-based batteries have a higher power density, a slightly greater energy density and the number of cycles is higher; many sealed construction types are available. From a technical point of view, NiCd batteries are a very successful battery product; in particular, these are the only batteries capable of performing well even at low temperatures in the range from -20°C to -40 °C. Large battery systems using vented NiCd batteries operate on a scale similar to lead acid batteries. Like lead, cadmium is a hazardous material. NiMH batteries have much higher energy densities (weight for weight). NiMH batteries are far safer than lithium ion batteries.

14.3.14.3 Lithium ion (Li-ion, LiPoly)

High cell voltage levels of up to 3.7 nominal volts mean that the number of cells in series with the associated connections and electronics can be reduced to obtain the target voltage. Another advantage of Li-ion batteries is their high gravimetric energy density, and the prospect of large cost reductions through mass production. Lithium ion batteries generally have a very high efficiency, typically in the range of 95% - 98%. Nearly any discharge time from seconds to weeks can be realized, which makes them a very flexible and universal storage technology. Standard cells with 5000 full cycles can be obtained on the market at short notice, but even higher cycle rates are possible after further development, mainly depending on the materials used for the electrodes.

Safety is a serious issue in lithium ion battery technology. Most of the metal oxide electrodes are thermally unstable and can decompose at elevated temperatures, releasing oxygen which can lead to a thermal runaway. To minimize this risk, lithium ion batteries are equipped with a monitoring unit to avoid over-charging and over discharging. Usually a voltage balance circuit is also installed to monitor the voltage level of each individual cell and prevent voltage deviations among them.

14.3.14.4 Metal air (Me-air)

A metal air electrochemical cell consists of the anode made from pure metal and the cathode connected to an inexhaustible supply of air. For the electrochemical reaction only the oxygen in the air is used.

14.3.14.5 Sodium sulphur (NaS)

Sodium sulphur batteries consist of liquid (molten) sulphur at the positive electrode and liquid (molten) sodium at the negative electrode; the active materials are separated by a solid beta alumina ceramic electrolyte. The battery temperature is kept between 300 °C and 350 °C to keep the electrodes molten. NaS batteries reach typical life cycles of around 4500 cycles and have a discharge time of 6.0 hours to 7.2 hours. They are efficient (AC-based round-trip efficiency is about 75%) and have fast response.

14.3.14.6 Sodium nickel chloride (NaNiCl)

The sodium nickel chloride (NaNiCl) battery, better known as the ZEBRA (Zero Emission Battery Research) battery, is – like the NaS battery – a high-temperature (HT) battery. Its operating temperature is around 270°C, and it uses nickel chloride instead of sulphur for the positive electrode. NaNiCl batteries can withstand limited overcharge and discharge and have potentially better safety characteristics and a higher cell voltage than NaS batteries. They tend to develop low resistance when faults occur and this is why cell faults in serial connections only result in the loss of the voltage from one cell, instead of premature failure of the complete system.

14.3.15 Flow batteries

A flow battery is also a rechargeable battery, but the energy is stored in one or more electroactive species which are dissolved in liquid electrolytes. The electrolytes are stored externally in tanks and pumped through the electrochemical cell that converts chemical energy directly to electricity and vice versa. The power of a flow battery is defined by the size and design of the electrochemical cell, whereas the energy depends on the size of the electrolyte storage tank(s). Flow batteries can be fitted to a wide range of stationary applications. Flow batteries are classified into redox flow batteries and hybrid flow batteries.

14.3.15.1 Redox flow

In redox flow batteries (RFB) two liquid electrolyte dissolutions containing dissolved metal ions as active masses are pumped to the opposite sides of the electrochemical cell. The electrolytes at the negative electrode are called 'analytes', and the electrolytes at the positive electrodes 'catholytes'. During charging and discharging the metal ions stay dissolved in the fluid electrolyte as liquid; no phase change of these active masses takes place. The anolyte and catholyte flow through porous electrodes, separated by a membrane which allows protons to pass through it for the electron transfer process. During the exchange of charge a current flows over the electrodes, which can be used by a battery powered device. During discharge the electrodes are continually supplied with the dissolved active masses from the tanks; once they are converted the resulting product is removed to the tank. Theoretically, a RFB can be "recharged" within a few minutes by pumping out the discharged electrolyte and replacing it with recharged electrolyte.

14.3.15.2 Hybrid flow

In a hybrid flow battery (HFB) one of the active masses is internally stored within the electrochemical cell, whereas the other remains in the liquid electrolyte and is stored externally in a tank. Therefore hybrid flow cells combine features of conventional secondary batteries and redox flow batteries: the capacity of the battery depends on the size of the electrochemical cell. Typical examples of a HFB are the Zn-Ce and the Zn-Br systems. In both cases the anolyte consists of an acid solution of Zn^{2+} ions. During charging Zn is deposited at the electrode and at discharging Zn^{2+} goes back into solution. As membrane a microporous polyolefin material is used; most of the electrodes are carbon-plastic composites.

14.3.16 Electrical storage systems

> **NOTE: Capacitance** *is a measure of ability to store electric charge.*

Electrical storage refers to the ability to store electric charge. Capacitance is the property that describes the storage of energy electrostatically (i.e., in an electric field). In other words, **capacitance** is a measure of ability to store electric charge. There are two closely related notions of capacitance, both of which are usually designated by the same term capacitance, and have the same SI unit of capacitance, the farad (F).

1. **Self capacitance** - Any object that can be electrically charged exhibits self capacitance. In a circuit, self capacitance is defined as the capacitive load, relative to circuit ground, that an electrode presents to the measurement system.
2. **Mutual capacitance** - the capacitive coupling between objects. The notion of mutual capacitance is particularly important for understanding the operations of the capacitor, one of the three fundamental electronic components (along with resistors and inductors).

A material with a large self capacitance holds more electric charge at a given voltage, than one with low capacitance.

- A 1 farad capacitor, when charged with 1 coulomb of electrical charge, has a potential difference of 1 volt between its plates.

In a mutual capacitance system, capacitance is a function only of the geometry of the design (e.g. area of the plates and the distance between them) and the permittivity of the dielectric material between the plates of the capacitor. For many dielectric materials, the permittivity and thus the capacitance, is independent of the potential difference between the conductors and the total charge on them.

14.3.16.1 Mutual capacitance

In a capacitor, the ratio of magnitude of charge on either conductor relates to the potential difference (voltage) between the conductors. Therein, for any given voltage (supplied by some power source), the amount of Q (charge that can be stored) increases with the amount of capacitance (which is the measure of the capacitor):

- Capacitance in farads = charge on either conductor / potential difference
- $C = Q / \Delta V$

The energy stored in a capacitor is found by integrating the value of work (W):

- $W = .5CV^2$

14.3.17 Capacitors

A capacitor (originally known as a 'condenser', and prior to that known as a permittor) is a passive two-terminal electrical component used to store energy electrostatically (i.e., in an electric field). Practical capacitors vary widely, but all contain at least two electrical conductors (plates) separated by a dielectric (i.e., insulator). In other words, a capacitor is (generally) two metal plates separated by an insulated material (i.e., non-conductive material). A capacitor can store electric energy when disconnected from its charging circuit, so it can be used like a temporary battery, or like other types of rechargeable energy storage systems. Conventional capacitors are commonly used in electronic devices, and the more recent supercapacitor technology has some ability to replace batteries.

Capacitor operation may be understood via analogies:

1. Beaty, W.J. (1996). *Capacitor complaints*. Amasci. [amasci.com]
2. Moffitt, B. (2014). *Capacitor pipe water flow analogy : DC*. Brandon Moffitt Channel. [youtube.be]

Capacitors store energy in an electrostatic field between their plates. Given a potential difference across the conductors (e.g., when a capacitor is attached across a battery), an electric field develops across the dielectric, causing positive charge (+Q) to collect on one plate and negative charge (-Q) to collect on the other plate, until the capacitor is fully "charged". The charges want to re-combine with one another, but they cannot because the insulator is preventing them from reaching each other. If a battery is attached to a capacitor, then no current will flow through the capacitor.

When capacitors are connected across a direct current DC supply voltage they become charged to the value of the applied voltage, acting like temporary storage devices and maintain or hold this charge indefinitely as long as the supply voltage is present. When a capacitor is connected to a circuit with a direct current (DC) voltage source, two processes, which are called "charging" and "discharging" the capacitor, will happen in specific conditions. By connecting the capacitor to the power supply, the charging phase occurs. Once the charging phase has finished, no more current flows through it. When the capacitor is disconnected from the power supply and connected to a load, then discharging occurs until the voltage between the capacitor's plates drops to zero, V_c=0.

When an alternating current is supplied to a capacitor, it will alternately charge and discharge at a rate determined by the frequency of the supply. Therein, the capacitance in AC circuits varies with frequency as the capacitor is being constantly charged and discharged, and a 'displacement current' will appear across the capacitor.

The greater the collection of charge on one surface of the capacitor, due to various parameters, such as surface area, the greater the energy capacitance of the capacitor.

14.3.17.1 Electrolytic capacitors

An electrolytic capacitor is a capacitor that uses an electrolyte (an ionic conducting liquid) as one of its plates to achieve a larger capacitance per unit volume than other types, but with performance disadvantages.

Electrolytic capacitor is the generic term for three different capacitor family members:

1. Aluminum electrolytic capacitors
2. Tantalum electrolytic capacitors
3. Niobium electrolytic capacitors

All electrolytic capacitors (e-caps) are polarized capacitors whose anode (+) is made of a particular metal on which an insulating oxide layer forms by anodization, acting as the dielectric of the electrolytic capacitor. A non-solid or solid electrolyte which covers the surface of the oxide layer in principle serves as the second electrode (cathode) (-) of the capacitor.

Like other conventional capacitors, electrolytic capacitors store the electric energy statically by charge separation in an electric field in the dielectric oxide layer between two electrodes. The non-solid or solid electrolyte in principle is the cathode, which thus forms the second electrode of the capacitor. This and the storage principle distinguish them from electrochemical capacitors or supercapacitors, in which the electrolyte generally is the ionic conductive connection between two electrodes and the storage occurs with statically double-layer capacitance and electrochemical pseudocapacitance.

14.3.17.2 Double-layer capacitors (DLC; supercapacitors)

NOTE: *Electrochemical capacitors go by a variety of names, including but not limited to: supercapacitor, super condenser, pseudocapacitor, electrochemical double layer capacitor, electric double layer capacitor, and ultracapacitor.*

Electrochemical double-layer capacitors (DLC) exist between classical capacitors used in electronics and general batteries. Electrochemical capacitors consist of two electrodes, a separator, electrolyte, two current collectors, and packaging. Within the electrochemical capacitor, charge is stored electrostatically, not chemically as in a battery. It has, as a dielectric, an electrolyte solvent, typically potassium hydroxide or sulfuric acid, and is actually two capacitors connected in series via the electrolyte. It is called a dual layer capacitor because of the dual layers within the structure, one at each electrode. The surface area is directly related to the amount of capacitance. The higher the surface area, the higher the capacitance of the capacitor.

There are two types of electrochemical capacitor technology, symmetric and asymmetric designs:

1. Symmetric designs are designs where both positive and negative electrodes are made of the same material with approximately the same mass, and which are available with aqueous or organic electrolytes.
2. Asymmetric designs uses a different material for the two electrodes, with one of the electrodes having much higher capacity than the other. Currently, asymmetric designs can use aqueous or organic electrolytes.

There are significant differences in the characteristics and performance of the four types which leads to a wide variety of products with many different possible applications. The fourth type is not included in this table since the type has active research programs directed toward its development, but it is currently not available as a commercial product.

The two main features are the extremely high capacitance values, of the order of many thousand farads, and the possibility of very fast charges and discharges due to extraordinarily low inner resistance which are features not available with conventional batteries.

Still other advantages are durability, high reliability, no maintenance, long lifetime and operation over a wide temperature range and in diverse environments (hot, cold and moist). The lifetime reaches one million cycles (or ten years of operation) without any degradation, except for the solvent used in the capacitors whose disadvantage is that it deteriorates in 5 or 6 years irrespective of the number of cycles. They are environmentally friendly and easily recycled or neutralized. The efficiency is typically around 90 % and discharge times are in the range of seconds to hours.

They can reach a specific power density which is about ten times higher than that of conventional batteries (only very-high-power lithium batteries can reach nearly the same specific power density), but their specific energy density is about ten times lower. Because of their properties, DLCs are suited especially to applications with a large number of short charge/discharge cycles, where their high performance characteristics can be used. DLCs are not suitable for the storage of energy over longer periods of time, because of their high self-discharge rate, their low energy density.

14.3.18 Superconducting magnetic energy storage (SMES)

Superconducting magnetic energy storage (SMES) systems work according to an electrodynamic principle. The energy is stored in the magnetic field created by the flow of direct current in a superconducting coil, which is kept below its superconducting critical temperature. 100 years ago at the discovery of superconductivity a temperature of about 4°K was needed. Much research and some luck has now produced superconducting materials with higher critical temperatures. Today materials are available which can function at around 100°K. The main component of this storage system is a coil made of superconducting material. Additional components include power conditioning equipment and a cryogenically cooled refrigeration system.

The main advantage of SMES is the very quick response time: the requested power is available almost instantaneously. Moreover the system is characterized by its high overall round-trip efficiency (85 - 90%) and the very high power output which can be provided for a short period of time. There are no moving parts in the main portion of SMES, but the overall reliability depends crucially on the refrigeration system. In principle the energy can be stored indefinitely as long as the cooling system is operational, but longer storage times are limited by the energy demand of the refrigeration system.

14.3.19 Thermal storage systems

Thermal (energy) storage (TES) systems store available heat by different means in an insulated repository for later use, including space heating or cooling, hot water production, and electricity generation. Thermal storage systems are deployed to overcome the mismatch between demand and supply of thermal energy, and thus, they are important for the integration of renewable energy sources. Second, utilization of waste heat in production processes by thermal energy storage reduces the final energy consumption (i.e., increases energy efficiency if the thermal energy is used or has the possibility of being used).

> **NOTE:** *Thermal energy is challenging to store, due to the ease of heat dissipation in physical systems.*

TES is applied in the field of power generation, production process heat, space heating and cooling, as well as the management of thermal energy processes in vehicles. These classifications of storage characteristics and applications result in specific operation parameters and designs of TES systems.

Thermal energy storage is achieved by different techniques, and can be subdivided into different technologies:

1. Storage of sensible heat - sensible
2. Storage of latent heat - phase change material
3. Thermo-chemical ad- and absorption storage - chemical sorption reaction

14.3.20 Sensible heat storage

The term ‘sensible heat’ indicates that the storage process can be *sensed* by a change of the temperature. The relation of the change in temperature and the stored heat is given by the heat capacity C_p. The storage of sensible heat is one of the best-known and most widespread technologies, with the domestic hot water tank as an example. The storage medium may be a liquid, such as water and solar ponds, thermo-oil, or a solid such as concrete or the ground. Thermal energy is stored solely through a change of temperature of the storage medium. Herein, the capacity of a sensible heat storage system is defined by:

1. The specific heat capacity, C_p.
2. The mass of the medium used.

Sensible heat can be stored in either solids and/or liquids:

1. *Solids:* metals, stones, salts, ceramics.
2. *Liquid:* water, thermal oil, molten salt.
3. *Liquid with solid filler material:* water with stones/ pebbles, oil with cast iron, molten salt with stone.

14.3.20.1 Solar Ponds

Water-based ponds may be used to capture the sun’s radiative energy. A solar pond is a pool of salt water that serves as a form of solar energy collection and sensible heat storage. A solar pond uses the principle of energy transfer by convection to heat saline water. This heated water solution may then be used for various purposes.

In general, a solar pond is a mass of shallow water about 1 or 2 metres deep with a large collection area, which acts as a heat trap. The pond contains dissolved salts to generate a stable density gradient. They are generally filled with saline water made with NaCl, $MgCl_2$, sodium carbonate, or sodium sulfate. Part of the incident solar radiation entering the pond surface is absorbed throughout the depth, and the remainder is absorbed at the very dark, black bottom. If the pond were initially filled with fresh water, the lower layers would heat up, expand and rise to the surface. Because of the relatively low conductivity, the water acts as an insulator and permits high working temperature (over 90˚C) to develop in the bottom layers. Hence, a gradient is maintained at varying densities. The bottom is the most dense and is used as a storage zone. Above the bottom layer is a non-convective zone, or insulation zone, with a density gradient which facilitates a temperature gradient as well. This layer functions as insulation. There is no convection in the gradient layer because even though the warm water would normally rise, the high salt concentration at lower levels does not allow the water to be light enough to float up as it warms. This prevents heat in the bottom from reaching the top of the pond. The top layer, or surface zone, is convective due to wind-induced mixing and daily heating and cooling. The hot brine, or salt water, on the bottom may be extracted and used for direct heating and low-temperature production uses like drying crops and agricultural shelter heating.

The problem with solar ponds is that it is essential to have a controlled saline density gradient, which is quite difficult to maintain. Additionally, the pond must be kept free of dirt and other light-absorbing materials. Thus, for large scale operations, the difficulties are too great to rely upon solar ponds for efficient heat production.

Here is one possible system for converting the heat energy from the salt water in the pond to electricity. The hot brine is pumped from the bottom of the solar pond through an evaporator (where it transfers heat to an organic 'working fluid,), and then, returned to the pond. The organic working fluid is heated in the evaporator, turns into a vapor, thereby producing sufficient pressure to spin a turbine connected to a generator. Therein, the vapor transfers some of its kinetic energy to the turbine. The cooler vapor is pumped to the condenser where it is condensed to a liquid as it transfers energy to the cold water being pumped through the tubes of the condenser. The organic liquid is now pumped to the evaporator to continue the process. As the gradient layer diffuses as time passes, new freshwater and salt water can be pumped into the pond to maintain a sufficient gradient layer.

14.3.20.2 Inter-seasonal thermal storage

Seasonal thermal energy storage (STES) allows heat or cold to be used months after it was collected from waste energy or natural sources. The material can be stored in contained aquifers, clusters of boreholes in geological substrates such as sand or crystalline bedrock, in lined pits filled with gravel and water, or water-filled mines. STES systems can be divided into:

1. Underground systems
2. Surface and above ground systems

14.3.21 Latent heat storage

In contrast to the storage of sensible heat, latent thermal energy storage (LHTES) cannot be sensed: The energy which is absorbed or released is stored by a *phase transition*, which takes place at a constant temperature, and therefore, appears to be latent. Materials used for latent heat storage are called PCMs (phase change materials), because the heat storage is achieved by a phase change of the storage medium. In other words, 'latent heat' is the energy exchanged during a phase change, such as the melting of ice. It is also called "hidden" heat, because there is no change of temperature during energy transfer.

Latent heat storage is accomplished by using phase change materials (PCMs) as storage media. PCMs include, but are not limited to the following, which are divided by in-/organic, and then, solid-solid or solid-liquid:

1. Organic phase change materials (organic PCMs)
   - *Solid-liquid:* paraffins
2. Inorganic phase change materials (inorganic PCMs)
   - *Solid-solid:* salt
   - *Solid-liquid:* water/ice, salt hydrates, salt/molten salt

The best known latent heat – or cold – storage method is the ice cooler (ice box), which uses ice (an inorganic PCM) in an insulated box or room to keep food cool during hot days.

Currently, most PCMs use the solid-liquid phase change, such as molten salts as a thermal storage medium or concentrated solar power (CSP) plants. The advantage of latent heat storage is its capacity to store large amounts of energy in a small volume and with a minimal temperature change, which allows efficient heat transfer.

14.3.21.1 Solar molten salt system

Molten salt is salt which is solid at standard temperature and pressure (STP) but enters the liquid phase due to elevated temperature. In this system, salt becomes molten once heated by a concentrated solar radiation system. It is then transported to a hot salt storage tank. To produce electricity, the hot salt passes through a steam generator that powers a steam turbine. Subsequently, the cold salt (still molten) is stored in a second tank before it is pumped to the solar tower again. The main disadvantages are the risk of liquid salt freezing at low temperatures, and the risk of salt decomposition at higher temperatures. In solar trough plants a dual-medium storage system with an intermediate oil/salt heat exchanger is preferred. Typical salt mixtures, such as $NaK\text{-}NO_3$, have freezing temperatures >200°C, and storage materials and containment require a higher volume than storage systems for solar tower plants.

14.3.22 Thermo-chemical heat storage

A thermochemical heat storage (TCS) system uses the enthalpy of a reaction ΔH. In reactions featuring a positive change of ΔH (endothermic reaction) heat can be stored. The energy can be released by a backward reaction ($\Delta H < 0$) afterwards. This chemical reaction always involves gas phase reaction:

1. *Solid - gas reaction:* dissociation reactions and adsorption processes

2. *Liquid - gas reaction:* absorption in alkaline or acid solution
3. *Gas - gas reaction:* methane reforming and ammonia dissociation

Sorption (adsorption, absorption) storage systems work as thermo-chemical heat pumps under vacuum conditions and have a more complex design than sensible or latent heat systems. Herein, heat from a high-temperature source heats up an adsorbent (e.g. silica gel or zeolite), and vapour (working fluid; e.g., water) is desorbed from this adsorbent and condensed in a condenser at low temperatures. The heat of condensation is then withdrawn from the system. The dried adsorbent and the separated working fluid can be stored as long as desired. During the discharging process the working fluid takes up low-temperature heat in an evaporator. Subsequently, the vapour of the working fluid adsorbs on the adsorbent and heat of adsorption is released at high temperatures. Depending on the adsorbent/working fluid pair the temperature level of the released heat can be up to 200°C and the energy density is up to three times higher than that of sensible heat storage with water.

Because of the possibility of storing the sorption compounds separately without the loss of energy, thermochemical storage is appropriate for thermal energy storage over large period of times.

14.4 Grid/network connectivity and power quality

Energy storage systems can be on-grid (network connected) or off-grid (network dis-connected).

14.5 Battery technology as energy storage

NOTE: *If electric current is like water, then in application, a battery is like a water pump. A pump takes in water at low pressure and does work on it, ejecting it at high pressure. A battery takes in charge at low voltage, does work on it and ejects it at high voltage. Batteries, however, do not store electric charges, and hence, they are not analogous to a water balloon shooting out water.*

A battery is a device that converts chemical power (energy) into electrical power (energy), and vice versa. Batteries supply electricity by producing voltage and delivering direct current (DC). Batteries do not produce AC voltage. Also, batteries do not store electricity, but rather store a series of chemicals, and through a chemical process electricity is produced. Hence, it is inaccurate to say that batteries store DC or DC voltage. Although a current can be described as moving electrical charges, it is not true that these charges are "stored in a battery". Batteries store chemical potential energy, which is released as DC voltage when connected to an electrical circuit. Batteries are an electrochemical storage carrier. The energy is stored chemically, but released as electricity. Through a chemical reaction process the battery creates and releases electricity as needed by the electrical system or devices. Batteries are sometimes considered electron pumps.

CLARIFICATION: *Electrical current is the movement of charged particles, such as electrons or ions, through a conductor.*

In composition, a battery is a technological device consisting of one or more electrochemical cells (voltaic cells). A voltaic cell is an electrochemical cell that uses a chemical reaction to produce electrical power (electrical energy). Simplistically, batteries contains atoms and molecules separated into ions that generate a voltage drop across their terminals.

CLARIFICATION: *A* **cell** *is the smallest, packaged form a battery can take and is generally on the order of one to six volts. A* **module** *consists of several cells generally connected in either series or parallel. A battery* **pack** *(battery bank) is then assembled by connecting modules together, again either in series or parallel. The term 'battery' and 'cell' are often used interchangeably; technically, however, a battery is made up from a group of cells.*

A battery releases energy at a more or less constant or flat voltage until depleted. A battery is used where a constant potential difference has to be maintained.

14.5.1 Battery components

All batteries have at least the following components, including two terminals (i.e., electrodes):

1. **The anode (terminal)** - an electrode where oxidation occurs (in a voltaic cell). This is generally a metal material.
2. **The cathode (terminal)** - an electrode where reduction occurs (in a voltaic cell). This is generally a different metal material than the anode material.
3. **The electrolyte (the ionic conductor)** - provides the medium for transfer of charge as ions inside the cell between the anode and cathode. The electrolyte is typically a solvent containing dissolved chemicals providing ionic conductivity. It should be a non-conductor of electrons to avoid self discharge of the cell.
4. **The separator** - electrically isolates the positive and negative electrodes. It is a permeable membrane placed between a battery's anode and cathode. The main function of a separator is to keep the two electrodes apart to prevent electrical short circuits while also allowing the transport of ionic charge carriers that are needed to close the circuit during

the passage of current in an electrochemical cell.

CLARIFICATION: *In a galvanic (voltaic) cell, the anode is considered negative (-ve) and the cathode is considered positive (+ve). This seems reasonable as the anode is the source of electrons, and the electrons flow to the cathode. However, in an electrolytic cell, the anode is positive (+ve), while the cathode is negative (-ve).*

14.5.1.1 Anodic index-galvanic corrosion (anodic index

The anodic index is a table showing the compatibility of different metals. This parameter is a measure of the electrochemical voltage that will be developed between two different metals. To find the relative voltage of a pair of metals it is only required to subtract their anodic indices. (Wheeler, 1972)

For example,

- The potential difference between iron and copper is approximately 0.4v.
- The potential difference between zinc and carbon is approximately 1.5v.

14.5.2 Battery operation

Batteries use a chemical reaction to do work on charge and produce a voltage between their output terminals. This voltage can be connected to an electrically conductive circuit (a load) to produce direct current (DC) electrical power. When a load completes the circuit between the two terminals, the battery produces electricity through a series of electromagnetic reactions between the anode, cathode, and electrolyte. Batteries operate based on the separation of [electric] charge in a chemical solution (i.e., an energy gradient), which produces an electromotive force (voltage) between their terminals. When the terminals are connected by means of an appropriate electrical conductor, then direct current will flow between the terminals powered by the pressure of the voltage (and, electrochemical discharge will occur).

Charging and discharging refer to the direction of current through a battery cell, and the type of chemical reaction that follows the current. Batteries are discharged or charged due to oxidation and reduction reactions therein.

1. **Charging** - the process of separating charge within a battery by providing DC electrical power to the battery.
2. **Discharging** - the process of releasing DC electrical power and reuniting charges via a conductive circuit.

Some batteries, due to their system's composition, cannot be recharged. Rechargeable batteries pump the charges back to their separate sides during charging, strengthening the electric field all over again. By reversing electrical current flow in a rechargeable battery, the chemical process is reversed, thus charging the battery. The cycle of discharging and charging is repeated continuously and is called **battery cycling**. All rechargeable batteries have a cycling lifespan (i.e., they can be cycled, discharged and recharged, a certain number of times before they need will no longer function without maintenance). Further, energy density decreases as a battery wears out.

NOTE: *A battery charger, or recharger, is a device used to put energy into a secondary cell (or rechargeable battery) by forcing an electric current through it.*

During discharge operation, the anode (terminal) experiences an oxidation reaction in which two or more ions (electrically charged atoms or molecules) from the electrolyte combine with the anode, producing a compound and releasing one or more electrons. In other words, the internal chemical reaction within the battery between the electrolyte and the negative metal electrode produces a build up of free electrons, each with a negative charge, at the battery's negative (-) terminal - the anode. At the same time, the cathode goes through a reduction reaction in which the cathode substance, ions and free electrons also combine to form compounds. In other words, the chemical reaction between the electrolyte and the positive (+) electrode inside the battery produces an excess of positive (+) ions (atoms that are missing electrons, thus with a net positive charge) at the positive (+) terminal - the cathode of the battery. The reaction in the anode produces a direct flow of electrons, and the reaction in the cathode absorbs them. The net product is electricity. The electrical (pump) pressure or potential difference between the + and - terminals is called voltage or electromotive force (EMF).The battery will continue to produce electricity until one or both of the electrodes run out of the substance necessary for the reactions to occur.

Batteries create electron flow in a circuit by exchanging electrons in ionic chemical reactions, and there is a limited number of molecules in any charged battery available to react, hence there is a limited amount of total electrons that any battery can propel through a circuit before its energy reserves are exhausted.

CLARIFICATION: *Reactants in a battery are separated internally by an electrolyte that provides for ion transfer, and externally by an electrical conductor (between the terminals), which provides for electron transfer. When a load is connected to the battery via the external conductive circuit, excess charges present in the negative ions of electrolytes on negative terminal (deposited during charging process) flow through the conductor, until they reach the positive terminal where they combine with positive* *ions, neutralizing the charge (the equilibrium condition). Under normal equilibrium conditions the electrical potential inside the battery exactly equals the chemical potential, and hence, there is no voltage and no electron flow. Normally, a battery which is not shorted out or connected to a load is under equilibrium conditions, meaning the chemical potential inside the battery exactly equals the electrical potential. Under these conditions, no charge carriers flow. If the positive terminal of the battery is connected to the negative terminal through some load, then the charge carrying electrons at the negative terminal of the battery will flow through the conductive load to the positive terminal.*

A battery is a chemical reactor where red-ox reaction happens. In the battery there is some medium, called an electrolyte, which can conduct only ions, but not electrons. Each voltaic cell consists of two half-cells connected in series by a conductive electrolyte containing anions and cations. One half-cell includes electrolyte and the negative electrode, the electrode to which anions (negatively charged ions) migrate; the other half-cell includes electrolyte and the positive electrode to which cations (positively charged ions) migrate. Redox reactions power the battery. Cations are reduced (electrons are added) at the cathode during charging, while anions are oxidized (electrons are removed) at the anode during charging. During discharge, the process is reversed. The electrodes do not touch each other, but are electrically connected by the electrolyte. Some cells use different electrolytes for each half-cell. A separator allows ions to flow between half-cells, but prevents mixing of the electrolytes. The reaction is driven forward by the chemical potential, an energy gradient. If the battery is not at 0 charge ("flat"), there will be a voltage between the electrodes.

INSIGHT: *When a battery is connected to a circuit, charges moves in the direction which diminishes the chemical potential energy in the battery.*

Because a battery is a system of separated charge, it has a static electric field. When a battery is powering an electrical load (i.e., discharging), over time, the separated charges get reunited via the conductive circuit, and the electric field between the positive and negative electrodes becomes weaker. Hence, as a battery is discharging (i.e., providing electrical power), its voltage (EMF) will drop over time. In a discharing battery, the chemical process is a decoupled redox reaction. As charges move around the circuit, the electrical potential inside the battery is reduced until equilibrium between the chemical potential and electrical potential is once again achieved.

Different metals have different affinities for electrons. When two dissimilar metals (or metal compounds) are put in contact or connected through a conducting medium there is a tendency for electrons to pass from the metal with the smaller affinity for electrons, which becomes positively charged, to the metal with the greater affinity which becomes negatively charged. A potential difference between the metals will therefore build up until it just balances the tendency of the electron transfer between the metals. At this point the 'equilibrium potential' is that which balances the difference between the propensity of the two metals to gain or lose electrons.

INSIGHT: *If you have two different metals in your body, particularly in your mouth, there will likely exist an electrical [galvanic] current between them. Such currents can and will interfere with the body's natural processes, potentially leading to states of dis-ease.*

14.5.3 Electrochemical cell types

Many types of electrochemical cells have been produced, with varying chemical processes and designs, including galvanic cells, electrolytic cells, fuel cells, flow cells and voltaic piles.

1. **Wet cell** - A wet cell battery has a liquid electrolyte. Other names are flooded cell, since the liquid covers all internal parts, or vented cell, since gases produced during operation can escape to the air.
2. **Dry cell** - A dry cell uses a paste electrolyte, with only enough moisture to allow current to flow. Unlike a wet cell, a dry cell can operate in any orientation without spilling, as it contains no free liquid, making it suitable for portable equipment. The common zinc–carbon battery (dry Leclanché cell) and zinc-manganese dioxide (alkaline cell) are both dry cell batteries.
3. **Molten salt** - Molten salt batteries are primary or secondary batteries that use a molten salt as electrolyte. They operate at high temperatures and must be well insulated to retain heat.
4. **Reserve battery** - A reserve battery can be stored unassembled (unactivated and supplying no power) for a long period (perhaps years). When the battery is needed, then it is assembled (e.g., by adding electrolyte); once assembled, the battery is charged and ready to work.

14.5.4 Battery condition parameters

This section describes some of the variables used to describe the present condition of a battery.

1. **State of Charge (SOC in %)** – An expression of the present battery capacity as a percentage of maximum capacity. SOC is generally calculated using current integration to determine the change in battery capacity over time.

2. **Depth of Discharge (DOD in %)** – The percentage of battery capacity that has been discharged expressed as a percentage of maximum capacity. A discharge to at least 80% DOD is referred to as a deep discharge.
3. **Terminal Voltage (V)** – The voltage between the battery terminals with load applied. Terminal voltage varies with SOC and discharge/charge current.
4. **Open-circuit voltage (V)** – The voltage between the battery terminals with no load applied. The open-circuit voltage depends on the battery state of charge, increasing with state of charge.
5. **Internal Resistance** – The resistance within the battery, generally different for charging and discharging, also dependent on the battery state of charge. As internal resistance increases, the battery efficiency decreases and thermal stability is reduced as more of the charging energy is converted into heat.

14.5.5 Battery energy and power units

Average power determines how long a battery lasts. Batteries have a limited amount of energy they can supply before they need to be replaced or recharged. The energy is typically measured in units of Joules. Batter capacity may be specified in Joules, or in amp-hours. If a 12 volt battery is rated at 100 amp-hours, then it means that - at least theoretically - it can supply 100 amps of current at 12 volts for one hour. Or, again theoretically, it can supply 1 amp of current at 12 volts for 100 hours. Either way, the total energy stored in the battery is S=3600VA Joules, where V is voltage in volts and A is the amp-hour rating, and 3600 is the number of seconds in an hour, and S is the number of Joules of energy the battery can supply. Power is measured in Watts, which is Joules/second. If a battery has S Joules of energy, and the average power you are drawing from it is P Watts, then the battery will be discharged in S/P seconds.

14.5.6 Battery technical specifications

This section explains the specifications you may see on battery technical specification sheets used to describe battery cells, modules, and packs.

> **CLARIFICATION:** *In describing batteries, discharge current is often expressed as a C-rate in order to normalize against battery capacity, which is often very different between batteries. A* **C-rate** *is a measure of the rate at which a battery is discharged relative to its maximum capacity. A 1C rate means that the discharge current will discharge the entire battery in 1 hour. For a battery with a capacity of 100 Amp-hrs, this equates to a discharge current of 100 Amps. A 5C rate for this battery would be 500 Amps, and a C/2 rate would be 50 Amps. Similarly, an* **E-rate** *describes the discharge power. A 1E rate is the discharge power to discharge the entire battery in 1 hour.*

1. **Nominal Voltage (V)** – The reported or reference voltage of the battery, also sometimes thought of as the “normal” voltage of the battery. All batteries will be damaged (if not explode and catch fire) if they are charged significantly above their nominal voltage.
2. **Cut-off Voltage** – The minimum allowable voltage. It is this voltage that generally defines the “empty” state of the battery. Some batteries will be damaged if they drop significantly below their cut-off voltage (e.g., lithium polymer batteries).
3. **Capacity or Nominal Capacity (Ah for a specific C-rate)** – A battery’s [electric current] capacity is the amount of electric charge it can deliver at the rated voltage. The coulometric capacity, the total Amp-hours available when the battery is discharged at a certain discharge current (specified as a C-rate) from 100 percent state-of-charge to the cut-off voltage. The amp-hour is a unit of battery energy capacity, equal to the amount of continuous current multiplied by the discharge time, that a battery can supply before exhausting its internal store of chemical energy. Therein, capacity is calculated by multiplying the discharge current (in Amps) by the discharge time (in hours), which decreases with increasing C-rate. For example, a battery rated at 100 A·h can deliver 5 A over a 20-hour period at room temperature. An amp-hour battery rating is only an approximation of the battery’s charge capacity, and should be trusted only at the current level, temperature and time specified by its technical specification.
   - Continuous current (amps) = amp-hour rating/ charge to discharge time (in hours)
   - Charge to discharge time (in hours) = amp-hour rating / continuous current (in amps)
4. **Energy or Nominal Energy (Wh - for a specific C-rate)** – The “energy capacity” of the battery, the total Watt-hours available when the battery is discharged at a certain discharge current (specified as a C-rate) from 100 percent state-of-charge to the cut-off voltage. Energy is calculated by multiplying the discharge power (in Watts) by the discharge time (in hours). Like capacity, energy decreases with increasing C-rate. A 12 volt battery rated for/ producing 100 amps of current has a 1200 watt-hour supply (12V · 100amp = 1200 W-h).
5. **Cycle Life (number for a specific DOD)** – The number of discharge-charge cycles the battery can experience before it fails to meet specific

performance criteria. Cycle life is estimated for specific charge and discharge conditions. The actual operating life of the battery is affected by the rate and depth of cycles and by other conditions such as temperature and humidity. The higher the DOD, the lower the cycle life.

6. **Specific Energy (Wh/kg)** – The nominal battery energy per unit mass, sometimes referred to as the gravimetric energy density. Specific energy is a characteristic of the battery chemistry and packaging. Along with the energy consumption of the vehicle, it determines the battery weight required to achieve a given electric range.
7. **Specific Power (W/kg)** – The maximum available power per unit mass. Specific power is a characteristic of the battery chemistry and packaging. It determines the battery weight required to achieve a given performance target.
8. **Energy Density (Wh/L)** – The nominal battery energy per unit volume, sometimes referred to as the volumetric energy density. Specific energy is a characteristic of the battery chemistry and packaging. Along with the energy consumption of the vehicle, it determines the battery size required to achieve a given electric range.
9. **Power Density (W/L)** – The maximum available power per unit volume. Specific power is a characteristic of the battery chemistry and packaging. It determines the battery size required to achieve a given performance target.
10. **Maximum Continuous Discharge Current** – The maximum current at which the battery can be discharged continuously. This limit is usually defined by the battery manufacturer in order to prevent excessive discharge rates that would damage the battery or reduce its capacity. Along with the maximum continuous power of the motor, this defines the top sustainable speed and acceleration of the vehicle.
11. **Maximum 30-sec Discharge Pulse Current** –The maximum current at which the battery can be discharged for pulses of up to 30 seconds. This limit is usually defined by the battery manufacturer in order to prevent excessive discharge rates that would damage the battery or reduce its capacity. Along with the peak power of the electric motor, this defines the acceleration performance (0-60 mph time) of the vehicle.
12. **Charge Voltage** – The voltage that the battery is charged to when charged to full capacity. Charging schemes generally consist of a constant current charging until the battery voltage reaching the charge voltage, then constant voltage charging, allowing the charge current to taper until it is very small. If a charger cannot detect when the battery is fully charged, then overcharging is likely. Overcharging will damage the battery, possibly causing it to catch fire and/or explode.
13. **Float Voltage** – The voltage at which the battery is maintained after being charge to 100 percent SOC to maintain that capacity by compensating for self-discharge of the battery.
14. **(Recommended) Charge Current** – The ideal current at which the battery is initially charged (to roughly 70 percent SOC) under constant charging scheme before transitioning into constant voltage charging.
15. **(Maximum) Internal Resistance** – The resistance within the battery, generally different for charging and discharging.

15 Energy demand requirements and usage monitoring

> **INSIGHT:** *After electrification, sleep became a disadvantage to feeding your family...because you could still be working for money.*

15.1 Reserve to production

The **reserves-to-production ratio** (**RPR** or **R/P**) is the remaining amount of a material resource, expressed in time. The reserve portion (numerator) of the ratio is the amount of a resource known to exist in an area and to be recoverable (proved reserves). The production portion (denominator) of the ratio is the amount of resource produced in one period (year) at the current rate.

1. RPR = (amount of known resource) / (amount used per year)
2. Units are time = amount / (amount/time)

This ratio is used to forecast the future availability of a resource to determine project life, and to determine whether more cultivation, harvesting, and/or exploration must be undertaken to ensure continued supply of the resource. Annual production of a resource can usually be calculated to quite an accurate number. However, reserve quantities can only be estimated to varying degrees of accuracy, depending on the availability of information and on the methods used to evaluate them.

Note that reserve and production rates are dynamic (constantly changing), and hence, this isn't a static calculation/analysis, but a dynamically recorded calculation. Also, reserves and production rates are not independent.

> **Renewable energy sources:** *energy sources that are capable of being renewed (in a sufficient timeframe), and for which the demand rate (or use rate) is less than or equal to the production rate.*

15.2 Gross and process energy requirements

> *Note: The* **embodied energy** *of a material refers to the amount of energy consumed in providing/producing that material. It is the energy consumed up to the end of the manufacturing process (cradle to gate). However, it may include delivery to the manufacturing site (cradle to site) or even the manufacturing processes into the completed product. This measurement can be used in comparing different materials.*

Technically, all energy inputs and energy outputs can be entered into a matrix, and have analyses run for any type of energy inquiry (i.e., the data can be parsed and analysed). Two of the most common analyses are known as 'gross energy requirement' and 'process energy requirement'.

The energy used in productive systems is measured based upon:

1. Gross energy requirement (GER) - the total amount of energy required for a system, project, object, service, or material. A measure of all the energy inputs that went into its creation and/or sustainment. The total energy consumed.
2. Process energy requirement (PER) - the energy consumed in the process directly undergone by building/creating a specified product (object, service, or material). A measure of the energy directly consumed in the process of producing the product, but not that of the second and higher generation consumption of the facility, transportation, etc.

15.3 Electrical energy demand

> **QUESTION:** *How do we know how much electric charge flow (i.e., electricity) to produce?*

Electrical generation and distribution equipment are designed and deployed to meet the maximum demand that all users/loads may require at one time. Hence, peak demand measurements are used to properly size the electric service ensure that there is sufficient generating capacity available (at all times).

1. The term kilowatt-hour (kWh) refers to the quantity of energy used; energy is signified by kWh. The kWh is a 'unit of energy'. Electrical energy actually used.
2. The term kilowatt (kW) refers to the electric power or rate (a.k.a., pace) at which this energy is used; demand is signified by kW. The kW is a 'unit of power'. The rate at which a load transforms electrical energy. Electrical energy moment by moment demand by a device.

The kilowatt hour (kWh, kW. h, kW-h) is how most home energy usage costs are calculated. The kWh measures [electrical] energy usage - equivalent to one kilowatt (1 kW) of power sustained for one hour. How much "energy" is in something or how much energy is used over a certain period of time.

- Watts * Time Used (in hours) = Wh
- Watts * Time Used / 1000 = kWh

Generator (if, then)

- If a device is rated to generate 1 kW of power (and it operates for one hour at that level), then it sustains

1 kWh of energy.

Load (if, then)

- If a device uses 100 watts over 10 hours, then it would utilize 1kWh of energy (100x10=1,000=1kW).
- For example, if a 40-watt bulb is used for 25 hours (40x25=1,000 watts = 1 kW), then it would use 1 kWh (at the 25 hour mark).

15.3.1 Load and supply

The balancing of load and supply on an electric power system is usually centrally controlled by a system operator or "Dispatch Centre". The dispatch centre continually how much generation is needed at each instant of time and issues orders (dispatch) to individual generating stations in merit order to supply the demand. The dispatch centre has a pretty good idea from forecasts of what will be required but they monitor frequency on an ongoing basis to ensure the balance is maintained.

Very rapid changes are dealt with automatically by stations which have some local frequency control. These stations respond very rapidly to changes in frequency by increasing power generated if frequency begins to drop. Those stations will need to keep some of their generating capacity in reserve for emergencies and the often receive payment just for being available to provide such a reserve.

Power generated must exactly equal power consumed. You have a few seconds to balance it, but frequency will start to rise or fall if there is any imbalance. However, voltage is not dependent on active power generated or consumed, frequency is.

A mentioned above the frequency (electrical cycles per second) is tied to the rotational speed of the generators (mechanical revolutions per second). In the USA the frequency is nominally 60 Hz but it varies slightly around this perhaps (59.97Hz to 60.03Hz). When extra power is drawn from the grid the generators feel the strain and slow down a bit - so the frequency slows down by a fraction of a Hz. This is the signal to put in more gas or steam and increase the power to your generators.

In the past control rooms that actually had big charts on the walls and dials that people turned. Nowadays the whole system is computerised. The longer term decisions still have human oversight but short term decision have to be made automatically because there isn't enough time for human intervention. In fact the fastest responding element is a speed controller on some of the generators (called a governor) which acts like cruise control on your car. When demand for electricity exceeds generation the generators start to slow down and the governors kick in and increase the gas.

To control the spin rate of the generators the system/technicials will, for example: apply more coal; turn on another turbine; in hydroelectric dams, open another gate and start another dynamo; in nuclear plants, pull out a couple of rods.

Some generators could alter their MW output, and some couldn't.

The frequency must stay the same as well. The frequency of the generator(revolutions per minute) goes up, as a result, the energy frequency at the generator output (Hz) goes up. Then it goes through a regulation circuit to synchronize it with the grid, to match the 50Hz(or 60Hz, if that's your thing).

15.3.2 Demand in a DC system

No content here yet.

15.3.3 Demand in an AC system

In an AC system, voltage doesn't vary with demand - frequency does.

1. If generation equals demand, then frequency will stay constant.
2. If generation is less than demand, then frequency goes down.
3. If generation exceeds demand frequency goes up.

15.4 Manufactured product energy usage label

The wattage listed is the maximum power drawn by the appliance. Every device, whether it is printed on the device itself or the transformer that powers the device, should have a label indicating the power consumption in Watts. The Watt number on the device may represent the device's maximum power rating, and not be representative of the amount of energy it is actually using when running. For instance, a computer's power supply may have a maximum rating of 400 watts, but someone uses the device at lower than 400 watts 100% of the time.

Take note, that in general, it is not possible to accurately calculate real world energy usage based on a manufacturer's label.

15.5 Market-based billing

> *Note: In general, electric meters record consumption (kWh).*

Typically, consumers in the market are charged for their electricity in terms of energy: the more energy used, the more they pay. A high-power appliance uses energy more rapidly than a low-power one, and therefore, costs more to operate (i.e., "run").

Your energy charges are based only on the total amount of energy you consume. Your demand charges are based on the highest level of electricity supplied at one time during the billing period and at the time of day it's needed by your business. In the market, electrical power companies use an electrical meter connected to a building to determine the power used, from which the total cost is calculated among a number of additional variables;

1. Time of day - peak hours (more cost); off-peak (lower cost)
2. Seasonal differentiated - e.g., more people using AC because of heat, more energy usage (more cost)

16 Energy density and power density

> **IMPORTANT:** *Modern physics is still confused over the concepts herein as is visible in the circularity of their definitions: (1) Energy is possessed by mass due to its motion and/or potential motion; (2) and, mass is a form of energy ($E=mc^2$).*

Energy density refers to the energy content/quantity of an energy carrier based on a mass or volumetric unit. Any material or energy resource (as a region of space) can be measured for its ability to release energy and do work. Similarly, any energy conversion system can be measured for its rate of transferring energy. These measures may be used as metrics for evaluating and comparing the performance of different energy technologies, materials and systems; or, they may be used as parameters for these design [specifications] for a new energy supply system.

> **CLARIFICATION:** *'Energy density' is sometimes known as 'work density'.*

If a system has a high 'potential energy density' or 'potential energy specificity', then it is able to contain ("store") a lot of energy in relatively little spatial area or mass. If a system has a 'high power density' or 'specific power', then per a given amount of time, it can output/ transfer relatively large amounts of energy based on its spatial area or mass. An energy resource with a lower 'energy density' or 'specific energy' means it takes more volume or mass to produce the same amount of work as a higher 'energy density' resource. Similarly, an energy system with a lower 'power density' or 'specific power' means it takes more time per volume or mass to transfer the same amount of energy as a higher 'power density' system.

The quantified presence of 'energy density' does not give information on how quickly this energy can be used/ transferred. This knowledge is contained in the term 'power density', which describes the rate at which energy can be put in and/ or transferred out. A high 'energy density' or 'energy specificity' does not necessarily mean a high 'power density' or 'power specificity'. Typically, having a high energy density means the presence of a low power density. In fact, in practical applications, a high energy density often equates to a low power density.

To better understand 'energy density' and 'power density' the case of a campfire could be used. To start the fire, kindling is used, because its high surface area-to volume ratio means that it burns quickly - a high 'power density'. However, once the fire is stable, kindling is no longer an optimal fuel source, because it burns too quickly. Hence, the fuel source is switched to logs, because they have a high 'energy density' and will burn for a longer period of time.

Metaphors are a useful way of understanding power density. For instance, dumping water out of a mug has a high power density, because it is capable of emptying all its contents almost instantaneously. Comparatively, if a jug with a 2cm spout is upended, it would take awhile to release its contents, giving it a low power density. For example, a tiny capacitor may have the same power output as a large battery, but because the capacitor is smaller, it has a higher power density.

> **NOTE:** *Magnetic energy density is the density of energy conveyed to the part of space occupied by the magnetic field.*

Just as energy can be separated at a top-level into potential energy and kinetic energy, energy density can be separated into potential energy density and kinetic energy density. Potential and kinetic energy can be released and transferred in the following ways:

1. Potential energy can be *released* from a system by a reaction (as 'energy density' or 'specific energy'), and *transferred* at a specific rate (as 'power density' or 'specific power').
2. Kinetic energy can be *extracted and/or released* from a flowing material by a device/system (as 'energy density' or 'specific energy'), and *transferred* at a specific rate (as 'power density' or 'specific power').

16.1 Energy [release] in relation to spatial region

Energy in a given region of space may be classified as potential or kinetic. Whether potential or kinetic, there is a maximum available energy per unit of space (volume or mass) that can be released or extracted. Therein, there are two calculated classifications for the amount of *available* energy within said region of space (measured as volume or mass), which create a total of four classifications:

1. [Potential or kinetic] 'energy density' - volume is the unit-region of space.
   A. Potential energy density
   B. Kinetic energy density
2. [Potential or kinetic] 'specific energy' - mass is the unit-region of space.
   A. Potential specific energy
   B. Kinetic specific energy

'Potential energy density' and 'potential specific energy' are applicable to a material potential energy resource, such as fuel or an energy storage medium (i.e., battery). 'Kinetic energy density' and 'kinetic specific energy' are applicable to a material kinetic energy resource, such as wind or flowing water.

'Energy density' is the amount of energy per unit weight (gravimetric energy density) or per unit volume (volumetric energy density). Note that the volume of a three dimensional space (i.e., spatial area) can be given in either metric cubic units (e.g., m^3) or liters (L). And, weight is otherwise known as mass. Energy density and specific energy are measures of the direct use/release of an energy source per spatial area unit as volume or mass. Energy density and specific energy answer the question: How much energy can be (or is being) stored in a volume/mass of space?

1. **Energy density (volumetric energy density)** - how much energy a system contains in relation to its volume. In other words, 'energy density' refers to how much work a given region of space (as a volumetric spatial unit) is capable of releasing (or exerting). It is typically expressed in watt-hours/volume or joules/volume: watt-hours per liter (Wh/L), joules per cubic centimeter (J/cm^3), joules per cubic meter (J/m^3), gigajoules per cubic meter (GJ/m^3), or megajoules per liter (MJ/L).
   - Potential energy density = energy output per volume of space derived as a result of a reaction (as chemical, thermal, nuclear fission or decay, electrochemical, or electrical).
   - Kinetic energy density = energy output per volume of space as a result of motion (as mechanical).
   - Energy density (Ed) is energy per volumetric spatial unit.
   - $E_d = E / V$
   - where, E is the energy released during utilization (e.g., combustion or energized circuit), and V is the volume of the fuel (as m^3 or liter).

2. **Gravimetric energy density (specific energy)** - how much energy a system contains in relation to its mass. In other words, 'specific energy' refers to how much work a given region of space (as mass unit) is capable of releasing (or exerting). It is typically expressed in watt-hours/mass or joules/mass: watt-hours per kilogram (Wh/kg), joules per gram (J/g), megajoules per kilogram (MJ/kg), gigajoules per ton (GJ/t).
   - Potential energy density = energy output per mass of space derived as the result of a reaction (as chemical, thermal, nuclear fission or decay, electrochemical, or electrical).
   - Kinetic energy density = energy output per mass of space as the result of motion (as mechanical).
   - Specific energy (Es) is energy per unit mass:
   - $E_s = E / m$
   - where, E is the energy released during utilization (e.g., combustion or energized circuit), and m is the mass of the fuel (as kg).

Energy density in an electric field is given by:

- Energy density = potential energy / volume = $1/2\epsilon_0 E^2$

Electric and magnetic fields "store" potential energy. In a vacuum, the (volumetric) energy density (in SI units) is given by:

- $U = (\varepsilon_0/2)E^2 + (1/2\mu_0)B^2$
- where, E is the electric field and B is the magnetic field. And, U is expressed in Joules per cubic meter (J/m^3).

In normal (linear and nondispersive) substances, the energy density (in SI units) is:

- U = 1/2 (E • D + H • B)
- where, D is the electric displacement field and H is the magnetizing field.

There are many different types of potential energy contained ("stored") in materials, and it takes a particular type of reaction to release each type of energy. A material (substance or system) can release [potential] energy in four types of reactions: nuclear; chemical; electrochemical; and electrical. In other words, these types of reactions are capable of releasing [potential] energy from a material.

Generally, only the useful (or releasable/extractable energy) is measured/quantified. For instance, in concern to chemical energy, chemically inaccessible energy is not applicable when accounting and calculating for energy density.

> **NOTE:** '*Energy density' (how much energy a region of space carries) does not provide sufficient information about 'energy conversion efficiency' (net output per input), or about 'embodied energy' (what the energy output requires in terms of harvesting, refining, distributing, and dealing with pollution, which all use energy themselves). However, the 'energy density' of a system can be calculated relative the inclusion or exclusion of [external] components required to express that energy (e.g., oxidisers, heat sink, temperature, and/or the energy output interface).*

Chemical reactions take in energy to break bonds and give off energy when they make bonds. Relatively large organic molecules like those of hydrocarbons have lots of weak carbon-carbon and carbon-hydrogen bonds which don't take a lot of energy to break. But when these molecules redox ("burn"), then the combustion products make lots of strong carbon-oxygen and hydrogen-oxygen bonds that give out a lot of energy when they form. However, a liquid fuel, for example, is only on chemical combination with Oxygen, that energy is released (converted) to heat. The output energy is more dependent on the bonds that form to make the products, than the bonds in the initial fuel.

> **CLARIFICATION:** *The amount of thermal energy released in a chemical reaction can be calculated through thermodynamic equations. The 'heat of combustion' is the total energy released as thermal energy (heat) when a substance undergoes complete combustion with oxygen under standard conditions (this information is often presented in units of kj/mol). Therein, the amount of energy released by the combustion of a given fuel is the result of the subtracting the energy required to break the bonds of the reactants from the energy released by the formation of bonds in the products. One measure of the chemical energy of a fuel is the 'heat of combustion'.*

When chemical reactions take place and bonds form, break, and reshape, the atom nuclei don't change in any way. What happens is that the electrons jump between atoms or groups of atoms and change orbits. The different orbits are bound by different energies and you can start with a configuration of atoms and electrons and end up with another configuration but the total energy stays the same. This is really all there is to it, and thinking about that will give you a lot of intuition in particular with "energetic" chemistry. The potential energy in the bonds becomes kinetic [thermal] energy.

> **NOTE:** *The energy density of a fuel (e.g., hydrocarbon) will vary depending upon its molecular makeup.*

Of note, energy per unit volume has the same physical units as pressure, and in many circumstances, it is a synonym for pressure. One could go so far as to say that energy, in every form, is simply pressure mediation. For example, the energy density of a magnetic field may be expressed as (and behaves as) a physical pressure. Similarly, the energy required to compress a compressed gas a little more may be determined by multiplying the difference between the gas pressure and the external pressure by the change in volume. Hence, pressure [in a fluid] may be considered to be a measure of energy per unit volume (energy density).

Using a battery as an example, 'energy density' is a measure of how much energy the battery can store, in a given size or mass, a characteristic of the battery chemistry and its packaging. A battery with a higher 'energy density' can power a load for longer than one with a lower 'energy density' and the same physical size or mass (and composition). Its units are in Wh/kg or Wh/cm^3. Note the use of hours in the unit: power x time = energy. Also, it has to be stressed that the calculated 'energy density' is related to a given discharge rate, temperature, battery size, average discharge voltage, and cut-off voltage. A battery can, for instance, have a higher 'energy density' when discharged at a lower discharge rate or to a final lower cutoff voltage:

1. Energy density of a battery = ((drain in amperes x service hours = capacity in amp-hours) x average

discharge voltage) / volume of battery in liters = Watt-hours/liter
2. Specific energy of a battery = ((current amperes x service hours = capacity in amp-hours) x average discharge voltage)/weight of battery in kilogram = Watt-hours/kilogram

16.2 Rate of energy transfer [power] in relation to spatial region

> **NOTE:** *Power is not sub-classified as potential or kinetic. Power is the rate of energy transfer. Power is measured in units: Watts = joules/ seconds*

Having a measure for 'energy density' or 'specific energy' does not give information on how quickly the energy can be used/transferred. This information is contained in the measure of 'power density' and 'specific power', which describes the rate at which energy can be released and/ or transferred. Hence, whereas 'energy density' refers to the capacity to do work, 'power density' refers to the speed at which work can be done. 'Power density' and 'specific power' are measures of power output as a time rate of energy transfer per spatial unit (volume or mass). Power density answers the question: How fast can a volume/mass of space deliver energy? There are two calculated classifications for the rate at which energy can transfer between regions of space (measured as volume or mass):

1. 'Power density' - volume is the unit-region of space.
   - Power density (W/cm^3 or W/m^3) = power / volume
2. 'Specific power' - mass as the unit-region of space.
   - Specific power (W/g or W/kg) = power / mass

'Power density' (a.k.a., volume power density) and 'specific power' (a.k.a., mass specific power) are the power correlates of 'energy density' and 'specific energy' -- the amount of power (time rate of energy transfer) between regions of space. 'Power density' and 'specific power' refer to the ability of a given system or region of space to deliver, or to take on, power. Power is transferred between regions of space via energy transforming systems (a.k.a., energy conversion devices), such as turbines, turbine-generators, batteries, fuel cells, motors, power supplies, combustion systems, and photovoltaic panels. Energy transformers convert a volume (volume power density) or mass (mass specific power) of energy at a specific rate.

> **NOTE:** *Since they release their energy quickly, in general, high power density systems can also recharge quickly.*

In terms of electrical energy, 'power density' (volume) and 'specific power' (mass) refer to the maximum current that can be drawn from a region of space (e.g., battery as volume and mass).

16.2.3.1 Power density and lasers

The power density of a laser beam can be determined by:

- Power density (W/cm^2) = 250/d^2 • power
- where, d is the diameter of the laser beam in millimeters. Assuming that the beam profile is uniform.

16.2.3.2 Power density and batteries

For a battery, 'power density' refers to the rate of energy release per unit of battery volume or weight (typically W/ dm3 or W/kg).

> **NOTE:** *A battery system can only supply the maximum power for a restricted period of (seconds or less) over the total amount of time it is supplying energy.*

16.2.3.3 Power density and radio

Power density is used by radio engineers to express power densities of isotropic antennas as a quotient of the transmitted power and the surface area of a sphere at a given distance (W/m^2).

16.2.3.4 Power density and heat engines

Power density of heat engines in kW/L (the second common use as volumetric power density of energy converters).

16.2.3.5 Power density and energy flux

A broader approach is to measure power density through the more universal measure of 'energy flux' as W/m^2 of horizontal area of land or water surface, rather than per unit of the working surface of a converter. In other words, power density is expressed as electricity generated per m2 of the area occupied (per period of time).

Wind power density, for example, refers to the energy flux across a turbine or to diffusion rates in fuel cells. Power density has been used recently in this sense in order to calculate a flux across the (vertical) area swept by a wind turbine (more on this in the wind power density section).

In terms of combustion to electrical energy, power density (as W/m^2) refers to the volume of raw material that would have to be extracted (relative to the material's specific energy) in order to meet the combustion system's desired production capacity (per period of time). The power density of the material could be calculated (per period of time), and then the power density of the entire system (including the technical combustion system and raw material) could be calculated.

In terms of a solar panel to electrical energy, power density (as W/m^2) refers to the power output of the panel relative to the surface area of the panel.

Subsequently, the power density of the entire arrayed system can be calculated. Note that the power density of solar installations must also account for space between panels, either for servicing in solar farms or for spacing between houses in rooftop solar installation.

In terms of a wind turbine, power density measures the flux of wind's kinetic energy moving through the working surface (the area swept by blades).

17 Energy and power safety

17.1 Warnings

No content here yet.

17.2 Failures

No content here yet.

17.3 Incidents types

Types of energy/power incidents include, but are not limited to:

1. All types of power:
   A. Fire
2. Electrical power:
   A. Shock and/or electrocution
3. Electromagnetic power:
   A. Electromagnetic burn
4. Mechanical, fluid, and combustion power:
   A. Suffocation
   B. Crushing
   C. Dismemberment
   D. Thermal burn
5. Nuclear power:
   A. Radiation burn (electromagnetic burn)

17.3.1 Native vs non-native electromagnetic radiation

Native (i.e., "natural") electromagnetic radiative energy is all around us. The three primary source of native EMR for us on earth are: 1) the earth; 2) the sun; and 3) the cosmos. All life on earth needs this native electromagnetic radiation (or at least a portion of it), and cannot function without it. In addition to native EMR, humankind has begun engineering electromagnetic radiation (i.e., non-native EMR or EMF). Biological organisms on earth have learned how to use EM radiation from the sun, and our species has learned how to use electromagnetic processes from reality to expand our functioning (e.g., radio telecommunications).

Humans are salt water, and they present a low-resistance path for electrical current, which will preferentially redirect through their bodies rather than through other substances.

1. Skin effect, inductances and capacitance are negligible at 60hz.
2. Emergency de-energization

"System Emergency" means the condition in the Electricity System when, due to the occurrence of one or more incidents, a part or the whole of the Electricity System experiences excessive frequency deviations or transmission voltage deviations, and in the opinion of the System Operator circumstances exist such that: (a) the safety of the Transmission System is at risk; (b) the reliable transmission of electricity is at risk; or (c) there exists a danger to life or property as a consequence of (a) or (b).

> **NOTE:** *The effects of EMR upon biological systems (and also to many other chemical systems, under standard conditions) depend both upon the radiation's power and its frequency.*

Hazardous energy is defined by the Canadian Standards Association (CSA) as: "any electrical, mechanical, pneumatic, chemical, nuclear, thermal, gravitational, or other energy that can harm people" (CSA Z460 "Control of Hazardous Energy - Lockout and Other Methods"). Some energy sources are obvious, such as electricity, heat in a furnace, or something that might fall. Others may be hidden hazards such as air pressure in a system or a tightly wound spring.

Not properly assessing and dissipating stored energy is one of the most common causes for workplace incidents that involve hazardous energy. Control of hazardous energy includes isolating the system from its primary power source and residual energy.

1. Hydraulic potential energy is the energy stored within a pressurized liquid. When under pressure, the fluid can be used to move heavy objects, machinery, or equipment. Examples include: automotive car lifts, injection moulding machines, power presses, and the braking system in cars. When hydraulic energy is released in an uncontrolled manner, individuals may be crushed or struck by moving machinery, equipment or other items.
2. Pneumatic potential energy is the energy stored within pressurized air. Like hydraulic energy, when under pressure, air can be used to move heavy objects and power equipment. Examples include spraying devices, power washers, or machinery. When pneumatic energy is released in an uncontrolled manner, individuals may be crushed or struck by moving machinery, equipment or other items.
3. Chemical energy is the energy released when a substance undergoes a chemical reaction. The energy is normally released as heat, but could be released in other forms, such as pressure. A common result of a hazardous chemical reaction is fire or explosion.
4. Radiation energy is energy from electromagnetic sources. This energy covers all radiation from visible light, lasers, microwave, infra red, ultraviolet, and X-rays. Radiation energy can cause health effects ranging from skin and eye damage (lasers and UV light) to cancer (X-rays).
5. Gravitational potential energy is the energy related to the mass of an object and its distance from the earth (or ground). The heavier an object is, and the further it is from the ground, the greater its gravitational potential energy. For example, a 1 kilogram (kg) weight held 2 metres above the ground will have greater gravitational potential energy then a 1 kg held 1 metre above the ground.
6. Mechanical energy is the energy contained in an item under tension. For instance, a spring that is compressed or coiled will have stored energy which will be released in the form of movement when the spring expands. The release of mechanical energy may result in an individual being crushed or struck by the object.

In most cases, equipment or systems will have safety devices built in. These safety devices include barrier guards and safeguarding devices to help protect workers during normal operations. However, during maintenance or repairs, these devices may have to be removed or by-passed. In these situations, a hazardous energy control program is needed.

A hazardous energy control program is used to maintain worker safety by preventing:

1. Unintended release of stored energy.
2. Unintended start-up.
3. Unintended motion.
4. Contact with a hazard when guards are removed or safety devices have been by-passed or removed.

Lockout is generally viewed as the most reliable way to protect an individual from hazardous energy because you are bringing the system to a zero energy state. When a system is in a zero energy state the hazard has been eliminated; thus, no hazardous energy exists. However, in some cases, using lockout is not practical because of its impact on operations and various other functions. Therefore, other controls can be implemented as long as adequate risk reduction of the hazard is obtained. This type of control means following a full set of steps to determine the hazards and risks of each task being performed, and determining what controls can be used to minimize and reduce risk to an adequate level. If an adequate level of risk cannot be achieved, then lockout will be the default method of control.

A voltage applied to a human body causes an electric current through the tissues, and although the relationship is non-linear, the greater the voltage, the greater the current. The threshold for perception varies with the supply frequency and with the path of the current, but is about 0.1 mA to 1 mA for mains-frequency electricity, though a current as low as a microamp can be detected as an electrovibration effect under certain conditions. If the current is sufficiently high, it will cause muscle contraction, fibrillation of the heart, and tissue burns. The lack of any visible sign that a conductor is electrified makes electricity a particular hazard. Death caused by an electric shock is referred to as **electrocution**.

17.3.2 Does it hurt more to be shocked by 110v or 240v AC?

One reason AC is more deadly is that any path which cause the current to pass through the body and cross the heart (i.e., left-hand-to-right-hand or hand-to-foot) will cause the heart to attempt to synchronize its beat to 60 Hz. The heart goes into fibrillation, and unless someone gets an AED on you within a couple minutes, that's the end. Heat is also generated by the flow of electrical current through body tissues, resulting in direct thermal injury (electrical burns) and possibly physical injury (entrance and exit wounds). In addition, the alternating current locks the muscles in a spasm, so you can't pull away. With DC, your greatest danger is thermal-physical injury at very high voltages, or just thermal injury at lower voltages. The reason DC feels much worse is that it causes the muscles to contract abruptly (whereas AC causes them to lock), so the physical effect is more painful. AC will freeze your muscles while DC will contract them. So if you hold a wire in your hand, DC will make you hold it even stronger, while AC will just "freeze" your muscles.

A higher voltage breaks down a poor insulator (e.g., the thin layer of non-conductive dry skin that covers the body), and once that insulator breaks down, the inner layers of the skin, and the muscles, are highly conductive. 15 mA is the lethal dose. That is why ground fault circuit interrupters (GFCI) can help prevent electrocutions, and are set to trigger at a 5 mV differential.

> **"***I have not tried the experiment, but I have read that a 9 V battery connected to two sharp needles will, if the needles are stuck into the skin, be very painful."*

Higher voltages are more dangerous because they break down poor dielectrics faster. Remember, at all times, 15 mA across the heart is all it takes. Current pushed/fed through the human body by voltage causes the body to react unpleasantly. Current is measured in amps, A. In the energy storage technology known as a battery, 'capacity' indicates how much power the battery pack can hold and is indicated in milliamp hours (mAh). In other words, mAh describes the measurement of how much load or drain (measured in milliamps) can be put on the battery for 1 hour, at which time the battery will be fully discharged.

However, DC tends much more to arcing than AC at points where current-carrying contacts are separated. The danger of fire and burning is significantly higher with DC compared to AC. The reason: an AC arc will be more efficiently extinguished due to the zero-crossings of current, 100 to 240 times a second (50 / 60 Hz). DC has no zero crossings, so an arc will form even at low voltages, around 40 - 50 volts. 250 VDC can easily cause an arc of one or more inches in length, if current is interrupted. Hybrid mechanical/electronic circuit breakers have only recently been developed for these DC current contact connections.

Of note, DC tends much more to arcing than AC. Any mechanical DC current connector, relay -- basically a mechanical current interrupter -- is dangerous, because of the DC arc, which will happen during the opening or closing of any DC current contact. Hybrid mechanical/electronic circuit breakers have been developed for these connections.

18 Power Symbols

A power symbol is a symbol indicating that a control activates or deactivates a particular powering or powered device. Universal power symbols are described in the International Electrotechnical Commission 60417 standard, Graphical symbols for use on equipment, appearing in the 1973 edition of the document (as IEC 417) and informally used earlier.

The well known on/off power symbol was the result of the logical evolution in user interface design. Originally, most early power controls consisted of switches that were toggled between two states demarcated by the words On and Off. As technology became more ubiquitous, these English words were replaced by the universal numeral symbols 1 and 0 (typically without serifs) to bypass language barriers. This "1" and "0" standard is still used on toggle power switches.

To create the symbol for a single on/off button, the "1" and "0" symbols were super-imposed onto each other to create the universally recognized power symbol used today. Because of widespread use of this symbol, a campaign was launched to add the set of characters to Unicode. February 2015, the proposal was accepted by Unicode and as of late 2015 the IEC power symbol family was in Stage 7 of Unicode character development and either in ISO approval ballot or pending ISO publication.

19 Energy and power fundamentals

Energy is a quantity that a given substance or system state/dynamic contains. An energy system is a sequence of energy calculations and energy transfers ("conversions"). The components of an energy (power) system must work together to transfer energy and [transform] power in such a way as to provide for higher functioning and the fulfillment of energy/power requirements.

> **NOTE:** *If energy were defined as the ability of a system to cause external action, then such action becomes sensible by force (displacement), heat (temperature), and light (EM radiation).*

In any given service system, energy has the following characterizations:

1. Energy is *expressed/experienced* as: effect, change, action, movement (or motion), behavior.
2. Energy is *carried/possessed* by: substances and systems.
3. Energy is *transferred* as: work or heat.
4. Energy is *required* for all: structures, processes, operations, functions.

Energy systems involve [at least] the following process stages/phases:

1. **Harvesting, harnessing, and/or collecting** natural energy sources/carriers.
2. **Transferring/transceiving** between different carriers of energy in order to store the energy or produce power. These transferring/transceiving processes are sometimes, and inaccurately, referred to as energy "transformation" and/or "conversion" processes. In actuality, the energy does not change form, it simply transfers carriers.
3. The **distribution** of energy carriers to their end-use application.
4. The **utilization** ("consumption") of energy by a service application.

Energy is a unifying abstract quantitative property that may be assigned to any region, system, object or substance in space-time (existence) via calculation, indicating a change to space-time (existence). The value of the property 'energy' is derived via a measured formulaic (algebraic) expression relative to the region of space (system) under consideration and the theorized principles-attributes that make up existence. The quantized value that is 'energy' indicates the relative degree to which motion is present and/or change of motion is possible in a given area of space-time. The formulas for deriving energy are human inventions -- the description of the natural law of energy conservation by means of a set of algebraic formulas is a human invention. Objects and systems have energy that can only be indirectly observed - by observing what the object or system does.

> **NOTE:** *Although 'energy' may be the most important concept in physics to define, it is also the most difficult to define, because it is axiomatic to the study of existence, which is not yet fully understood.*

Energy is not a thing; it is not matter and it cannot be reified. But, existence may be said to "carry" a property called 'energy'. 'Energy' is an abstract value of a system, and when calculated in "appropriate" ways, always turns out to be conserved.

> **INSIGHT:** *When energy is present, change is occurring or is possible.*

It is incorrect to conceive of energy as a substance-like entity that occupies space, but there are entities called "fields" which occupy space. There are electric and magnetic fields which occupy space, and are described by the "classical" electromagnetic theory of Clerk Maxwell and others, way back in the 19th Century. To conceptualise energy as having form or substance, you should be aware that such conceptualisations are not required by the mathematical structure of physics. There is nothing in the equations or laws which says that energy has either form or substance. It is nothing more than an abstract entity which can be calculated according to rules which it is the business of physics to discover.

> **NOTE:** *Magnetic fields are really electric fields under a Lorentz transformation. Electric fields do not occupy space in the sense that matter does. Instead, they occupy "counterspace". You cannot put two matter-objects (e.g., clay bricks) in one place, but you can superimpose two magnetic fields in the same space to get a total magnetic field (in "counterspace").*

Characteristics of a broad conception of energy:

1. Energy is associated with the action of organisms.
2. Physical systems and objects possess and expend energy.
3. Energy of itself does not cause anything.
4. Energy is associated with activity.
5. Energy is transferred/exchanged between carriers by processes.
6. Energy is a generalised kind of fuel associated with making life comfortable.
7. Energy is a kind of fluid which is transferred in some processes.

There is an inaccuracy in the following commonly seen definition of energy: **"**Energy is the ability to do work, where work is force times distance." The concept that energy is the ability to do work dates back to the seventeenth century and was put into question when energy was defined quantitatively as a conserved quantity in the 1840s. Within ten years, the enunciation of the second law of thermodynamics had shown this definition to be false. In thermodynamics, 'work' takes on a meaning which is broader than the "force times distance" concept of classical mechanics. Therein, work refers to either (1) a process of energy transfer, or (2) the energy being transferred. Not all energy can produce work -- according to the second law of thermodynamics there is always some part of the energy of a system (disordered) that cannot produce work. Hence, there does not seem to be much information content in the statement that "energy is the capacity to be energy that is being transferred (by some process or other)" or "energy is the ability to do force times distance".

> **INSIGHT:** *Spurred by the burgeoning industrial revolution, the theoretical physicists between 1840 and 1860 developed the "laws of thermodynamics". The first law (conservation of energy) told the industrialists that they could not get anything for nothing; the second law (increase in entropy) said that they could not even break even.*

Every time energy is transferred, some of its ability to do "work" is irretrievably lost. However, no such limitation applies to the conversion of work to heat; if a simplified definition of energy is needed, it might be described as the ability to produce heat (heating). While this definition is neither elegant nor useful, at least it is true. It is therefore misleading to leave heat out of any definition of energy. Thus, in every transfer of energy, there is energy available [given appropriate technology] to do "work", and energy not available to do "work" -- energy is the capacity to do 'work' and/or supply 'heat'. Therein, the specific expression/definition of 'work' is dependent upon the contextual medium possessing or transferring energy, and 'heat' is energy transfer between two objects of different temperatures:

- **Available energy**, that is energy which can, in principle, be transferred between systems by the process called "work".

Additionally, without the idea of "conservation", energy would mean nothing. Conservation of energy means that energy calculations, correctly performed, always balance to give a constant total. Note that the term 'conservation' does not mean the "saving of energy" or "not wasting energy".

The amount of work that can be obtained from energy depends on the degree of organization of the energy. Organization means all the molecules are moving in the same direction. If the same amount of energy is added to the random motion of the molecules, the result will be a rise in temperature.

Energy and mass are two different ways of expressing a certain property of a system, and that property is a conserved quantity. The equation $E = mc^2$ does not state that mass can be converted into energy. What it does say is that the total energy of a system can be found by multiplying its mass by a universal constant (or consistent). In a frame of reference affixed to the object, multiplying its mass by c^2 yields a quantity called 'rest mass energy'. All mass carries energy.

> **INSIGHT:** *Mass can only be determined when a particle is at rest. This is called 'rest mass'. Electromagnetic emissions cannot be put to rest, and therefore, the mass of these "particles" (if they even have mass) cannot be determined.*

Mass and energy are scalar quantities, while momentum is a vector quantity. Kinetic energy is scalar, it does not have direction. A scalar quantity has only magnitude, while a vector quantity has both magnitude and direction. Momentum is a vector quantity, because it has magnitude and direction. Although a scalar quantity may be separated into components, that doesn't make it a vector quantity. A vector is defined by how those components add up to a total. Vectors also follow specific coordinate transformation laws. As viewed from different coordinate systems, the magnitude of a vector would be the same, but the values of its components would be different. There are requirements for a physical quantity to be a vector: 1) it should have a direction; and 2) it should be added by laws of vector addition.

> **INSIGHT:** *We understand that matter is made of "atoms". We understand that atoms have "mass". We understand that there is a relationship between mass and energy. We understand that there is a relationship between vacuum and energy.*

At the center of all things with magnitude (e.g., magnetism), there is not that force (or inertia) modality. At the center of gravity there is zero gravity. At the center of magnetism there is zero magnetism. At the center of charge there is zero charge. Where there is not that force there is a "plane of inertia". There is no midsection to any magnet. Each new slice of a magnet will have its own plane of inertia. That tells you that there is no thing that is a plane of inertia. That tells you that the "block wall" or plane of inertia is not located there.

19.1 Energy and systems

The concept of energy can be applied to any existent system, wherein the following principles apply:

1. Energy is a derived quantity that may be assigned as a property of any existent system.
2. The total quantity of energy in a closed system is fixed.
3. The total quantity of energy in an open system is not fixed. In an open system, energy is brought into the system from the environment.
4. More 'energy' is required to transfer something across a system boundary, than within the system boundary.
5. Systems maintain themselves by cycling energy and matter. Ecosystems maintain themselves by cycling energy and material nutrients [obtained from external sources].

An 'energy process' is any change that a system undergoes from one equilibrium state to another. Therein, a 'path' is the series of states through which a system passes during a process. To describe a process completely, one should specify the initial and final states, as well as the path it follows, and the interactions with the surroundings.

19.1.1 Energy and thermodynamic systems

Thermodynamics does not have a direct etymological meaning, but it could be named the study of heat transfer. Anything in physics related to heat is classified [at least] as part of study of thermodynamics, and follows thermodynamic principles. There are four principles (or "laws") of thermodynamics (zero through three):

1. **The zeroth law of thermodynamics** - if a system A is in equilibrium with system B, and system B in turn is in equilibrium with system C, then systems A and C are in equilibrium with each other.
2. **The First Law of Thermodynamics** - the change in internal energy of a system is equal to the sum of the energy transferred to the system by "heat" and the "work" done on the system. This principle claims that energy is conserved.
3. **The Second Law of Thermodynamics** - the efficiency of heat engines must always be < 1.
4. **The Third Law of Thermodynamics** - the temperature of a system cannot reach absolute zero (0 K); as the system approaches absolute zero, entropy approaches a constant.

Thermodynamics divides the universe/reality into two parts: a system and its environment. Therein, there are three distinct types of thermodynamic system:

1. **Isolated systems** - no transfer/exchange of energy or matter with the environment.
2. **Closed systems** - transfer/exchange only energy, but not matter with the environment.
3. **Open systems** - transfer/exchange both energy and matter with the environment.

Thermodynamics concerns equilibrium states. Thermodynamic equilibrium is an axiomatic concept of thermodynamics. It is an internal state of a single thermodynamic system, or a relation between several thermodynamic systems connected by more or less permeable or impermeable boundaries. In non-equilibrium systems, by contrast, there are net flows of matter or energy. Equilibrium is a state of balance (no change). In an equilibrium state there are no unbalanced potentials (or driving forces) within the system.

1. **Thermal equilibrium** is if the temperature is the same throughout. The point at which heat transfer stops.
2. **Mechanical equilibrium** is if there is no change in pressure at any point of the system with time.
3. **Phase equilibrium** is if a system involves two phases and when the mass of each phase reaches an equilibrium level and stays there.
4. **Chemical/atomic equilibrium** is if the chemical/ atomic composition of a system does not change with time, that is, no chemical or atomic reactions occur.

The prefix iso- is often used to designate a process for which a particular property remains constant.

1. **Isothermal process** - a process during which the temperature remains constant.
2. **Isobaric process** - a process during which the pressure remains constant.
3. **Isochoric (or isometric) process** - a process during which the specific volume remains constant.
4. **Cycle** - A process during which the initial and final states are identical.

19.1.2 Thermodynamic energy flow types

A system and/or region of space can "gain" energy from its surroundings or "lose" energy to its surroundings. Technically, the system or region of space isn't actually gaining or losing some real thing called 'energy'; instead, the property, 'energy', which is assigned to that region of space or system is increasing or decreasing in value. As energy flows through systems it acquires different qualities:

1. **Enthalpy** - The amount of heat content used in a system at a constant pressure. In any system how much heat is used at the constant pressure termed as enthalpy. Enthalpy (h) is the measure of total energy content of a substance in a thermodynamic system. The SI unit for 'specific enthalpy' is joule per kilogram. It can be expressed in other specific quantities by $h = u + pv$, where u is the 'specific internal energy', p is the pressure, and v is specific volume, which is equal to 1ρ, where ρ is the density.
2. **Entropy** - Entropy is the measure of disorder, or a measure of "randomness". Entropy is a quantitative measure of the unavailability of a system's energy [to do work]. When a system receives an amount of heat (ΔQ), the system gains an entropy in the amount given by ΔQ/T, where T is the absolute temperature at which the heat transfer takes place. Entropy is a measure of the disorder of a thermodynamic system. Enthalpy is a measure of the total energy of a thermodynamic system. Entropy is a quantitative measure of the molecular disorder of a system. Entropy of the system increases when temperature gradients disappear or dissipate. Entropy of the system increases when concentration gradients disappear or dissipate. As energy is expended to do work, entropy decreases. If no energy is available in the system, its entropy level will remain constant or increase.
   - There are at least two kinds of entropies - thermodynamic and information-theoretic

ENERGY AS WORK AGAINST ENTROPY
If energy is the ability to do work, then without energy there is no ability to maintain structure against the entropic randomization of the universe -- without energy there is no ability to do anything. Here, we take high energy molecules or molecules in motion that ultimately derive their energy from the sun (i.e., the basic economy of the planet is energy from the sun, is photosynthesises) and transformationally redirect it into something we can use. In other words, we take that high potential environmental energy and convert it through a process of some kind into directed high potential [information] energy that allows our habitat (and bodies) to be powered. At a biological level this occurs through a quantum transducer known as 'mitochondria'. Biologically speaking, our organisms strip electrons off food in a similar, though significantly more complex, manner to the functioning of turbines in a hydroelectric dam through which water flows. When water goes through the turbines electrons are taken [by us] and fed into an electric grid through electron transmission. In fact, our mitochondria is a miniature example of this electron transport chain seen in hydroelectric or nuclear generating power systems. Our organisms take high energy [macro] nutrients as proteins, lipids, and carbohydrates and process them through an energy "powerhouse" we know as mitochondria to produce a set of high-energy intermediaries (e.g., atp, nadph) that are then directed and delivered to regions of the cell(s) that maintain function (i.e., muscles to contract or neurons to fire or digestive juices to be released or cells to replicate. This process is technically known as mitochondrial bioenergetics. At a practical level our diet and lifestyle play an important role in functionally maintaining the ability of these powerhouses to do that work. Fundamentally, matter is a form of coalesced energy (remember e = mc2), and molecules are essentially information rich "data packets" of energy.

entropies. Thermodynamic entropy should not be confused with the so-called 'information theoretic entropy' (a.k.a., intropy). 'Information theoretic entropy' is a measure of variety of message sources in communication systems.

3. **Exergy** - The generic name for the amount of work obtainable when some matter is brought to a state of [energetic] equilibrium with its surroundings by means of reversible processes. In thermodynamics, the exergy of a system is the maximum useful work possible during a process that brings the system into equilibrium with a heat reservoir. When the surroundings are the reservoir, exergy is the potential of a system to cause a change as it achieves equilibrium with its environment. Exergy is the energy that is available to be used. After the system and surroundings reach equilibrium, the exergy is zero The maximum useful work which can be obtained in a process in which system obtains dead state. For example, when a new source of energy, such as geothermal well, is discovered, the first thing the team does is estimate the amount of energy contained in the source. However, this information alone is of little value in deciding whether to build a power plant on the site. Additional information is needed in order to determine the work potential of the source; that is, the amount of energy that can be extracted as useful work. This property is called exergy. Note that destroyed exergy has been called anergy.
4. **Emergy** - is defined as the available energy of one kind required directly and indirectly to make a product or service. It is measured in emjoules (a unit referring to the available energy consumed in transformations). In energy systems, energy sources and components are connected with energy flows and arranged from left to right (generally), thus describe the order of increasing transformity. The transformity is defined as the emergy (in emjoules) of one kind of available energy required directly and indirectly (through all the pathways required) to make one joule of energy of another type. Transformity is the ratio of emergy to available energy.

19.2 The physics perspective on energy

INSIGHT: *Energy is a mathematical principle, not a description of a mechanism or anything physical.*

'Energy' is a physics term describing a mathematical principle - something that cannot be held or seen. Energy is simply a value that is algebraically calculated. Therein, energy is an abstract and quantized value (a property) that may be assigned to any given region of space-time, implying the presence of change or the possibility for change. The property, 'energy', is the result of mathematical expression and calculation, and cannot be reified. There is no such thing as "a form of energy", because energy does not have form - it is not an existent substance. Energy does not come in different forms. Hence, 'energy' is not capable of being transformed or converted. Instead, energy is an abstract, calculated value implying the transformation of existence, or potential for the transformation of existence, but it itself does not have existence or form. All systems in existence "carry" energy (i.e., have energy as a property assignable to their existence).

CLARIFICATION: *Energy is not a physical substance; it is not made of atoms. Matter is made of atoms. The distinctive names of energy, such as kinetic energy, potential energy, electrical energy, mechanical energy, thermal energy, arise because of the different systems by which energy is assigned as a property, not because there are different forms or types of energy. Energy is just energy. Different physical "storehouses" result in the different names, not different kinds of energy. Many papers cite the approach in the Feynman Lectures: there are a number of different physical quantities whose sum is always constant (as a physical law), so we call that sum "energy". These various physical quantities are typically called "forms of energy".*

The value of the property, 'energy', can be calculated for any region of space-time and/or system given the data available to complete the calculation. 'Energy' cannot be calculated in isolation, or measured directly. Conceptually, energy only exists in the context of two or more objects and the occurrence of a change (i.e., in the context of a system).

Energy is a universal property of every existent system that implies the calculated ability to change [given a cause in the real world]. And, all change in the universe implies the transfer of energy. A change in a system's state implies a transfer of energy, and a transfer of energy implies a change in a system's state. The presence of energy and its transfer underlies all change of existence in the universe. Its presence, as an expressed property, is a calculable necessity for the existence of movement, heat, electricity, and all life.

NOTE: *In physics, the presence of energy is required for change, but energy does not cause change. Differences in the real world (e.g., pressure and temperature gradients) cause change, and when change occurs, energy transfers (i.e., is recalculated). In other words, differences in pressure concentration and unbalanced forces cause change, not energy. But, when change occurs, energy is transferred. And, the quantity that is 'energy' can be calculated to determine the value required for an expressible change to occur. For instance,* *the energy requirement for lifting an object a set height can be calculated.*

Energy is not made of anything, energy is a term used to describe a trait of matter and non-matter fields. When matter has velocity, for example, it is said to have kinetic energy. Energy isn't 'stuff', it is a calculated quantity based on an equivalence relation; a property that systems have. It is a scalar value that is assigned to the state of a system, (i.e., energy is a tool or mathematical abstraction of a property of physical systems). This quantity is involved in (i.e., can be associated to) all processes of change.

Mathematically, energy is a conserved quantity, and hence, it is not possible to create new energy that is not already present (mathematically speaking) in the universe. Many different kinds of change take place in the real world, and for each kind of change the total energy before is calculated to equal the total energy after.

NOTE: *In common parlance, the terms "energy generation" and "energy utilisation" are confusing because, in fact, no energy can be created or destroyed.*

A classic textbook definition of energy is "the ability (or capacity) to do work". Even though this definition is now deprecated, it seems to have enduring power "Energy is the capacity to do work" is not only incomplete, but incorrect, because it ignores the 2nd Law of thermodynamics, which states that not all energy has the ability to do work. The 1st law says that energy is conserved, yet the 2nd law says that the ability to do work is not conserved, so this definition of energy leads to a logical contradiction. Further, although at a conceptual level the idea that energy is "the capacity to do work" makes sense for mechanical energy, but not for thermal and other "forms" of energy.

19.3 Energy classification contexts

The total energy of a system can be classified in a variety of ways. Hence, there are multiple "types" of energy, which depend upon the context of the classification. As a measure/property of the expression of something that exists, the presence of energy can be classified according to:

1. Spatial motion (of carrier/system)
2. Spatial length (of carrier/system)
3. Spatial medium (of carrier/system)
4. Pressure gradient (within carrier/system)
5. And, any combination thereof

19.3.1 Classified by: Spatial motion

Stating that an object "possesses" energy means that it is moving right now or has the ability to move due to its position [in a force field]. The part of the measure that is not currently represented as motion, but stays in the form of an ability to move due to its position is called, potential energy. While, the part corresponding to the motion right now is called kinetic energy.

In general, motion and the potential for motion is described as mechanical energy. Mechanical energy is the total energy an object has (when building up energy or when using it). The two known types of energy as per motion are: kinetic energy (current motion) and potential energy (potential for motion). Hence, the sum of potential energy and [macroscopic] kinetic energy is called mechanical energy.

Hence, from the perspective of spatial motion, energy can be classified in two fundamental ways:

1. As a flow: Energy <u>expressed</u> as motion is called **kinetic energy (mass and velocity)**. Kinetic energy is the energy an object "possesses" due to its motion. A force, driven by a cause, is required to accelerate a mass to its stated velocity. Kinetic energy is "gained" during acceleration and remains during motion. Kinetic energy represents explicit change.
2. As a potential for flow: Energy <u>associated</u> with position [in an electric or magnetic field] or condition is called **potential energy (position and condition)**. Potential energy is the energy associated with an object as a result of its position or condition inside a "force" field. Note here that potential energy is not stored energy. Energy can be stored in motion just as well as it can be stored in position. Potential energy represents a potential for future change (or motion). Energy due to position in space.

And yet, this arrangement of motion associated energy concepts could also be viewed from the following perspective classification:

1. How energy flows: electromagnetic energy, electrical energy, heat, work, etc.
2. How energy is stored: internal energy, kinetic energy (stored in motion), potential energy (stored in condition), energy density (stored in fuel), etc.

Simplistically, kinetic energy is energy "in motion" and potential energy is energy "at rest". Potential energy is often measured as positive or negative depending on whether they are greater or less than the energy of a specified base state or configuration, such as two interacting bodies being infinitely far apart.

When energy is present, then there ongoing change (i.e., movement) or the possibility of future change (i.e., future movement). That potential for future change can be "absorbed" by some sort of reservoir-like capability in existence (as potential energy), retained for a length of time (permittivity), then released again in the future as explicit change (as kinetic energy).

A existent\material system can have any combination of kinetic and potential energy, and both potential and kinetic energy can be "stored". Over time, energy is transferred from potential energy to kinetic energy, and then, back to potential energy constantly. This is referred to as the "conservation of energy" (or, "conservation of ether").

Potential energy only exists in the presence of a system. When a brick, for instance, is lifted, its potential energy is not increased. Instead, the potential energy of the system that consists of the two objects: the earth and the brick. Potential energy is not stored in either Earth or the brick, nor is it possible to apportion the potential energy between the two interacting objects. The first step to enlightenment is to learn not to speak or think about gravitational or electrical potential energy of electrons (or of any other kind of particle) and think instead of energy as a property of the whole system.

> **NOTE:** *Technology takes forms of potential energy and turns them into kinetic energy. Whenever energy is "used", it is kinetic energy.*

The material expression of kinetic and potential energy can take multiple forms. These forms represent a contextual measure (quantity) of the ability of a system or object to do work on another system or object in material reality. The interaction between energy and matter is described by its multiple forms of energy.

> **QUESTION:** *Energy is not a thing, it is a property of something; hence, how can it exist in different forms? Simply put, a "form" of energy is a mathematical expression, an equation that evaluates a quantity of energy in a given context. Thermodynamics is the study of any energy transformation that involves heat.*

Kinetic energy is calculated using the following formula:

- $E = 1/2mv^2$
- where, E is energy, measured in joules (J)
- where, m is mass, measured in kilograms (kg)
- where, v is velocity, measured in meters per second (m/s)

Notes about the expression:

1. The more mass a moving object has, the more kinetic energy it will possess at the same speed.
2. Because the velocity term in this formula is squared, velocity has a much larger effect than mass does on kinetic energy.
3. The larger an object (m) is and the faster it moves (v), the more kinetic energy (KE) it has.

Ways of harnessing macroscopic kinetic energy include, but are not limited to:

1. Wind power harnesses the kinetic energy possessed by moving bodies of air (wind) as they flow between atmospheric pressures.
2. Hydropower harnesses the kinetic energy of moving water as it moves (i.e., falls) in a "gravitational force" field.

Potential energy is calculated using the following formula:

- $E = gmh$
- where, g is the acceleration due to "gravity"
- where, h is height above inertial ground plane
- where, m is the mass of a body.

Notes about the expression:

- The larger an object is (m) and the more displaced it is (h) in a "force" field, the more potential energy (PE) it possesses.

19.3.2 Classified by: Spatial length

All energy types detectable in space-time may be classified by the length-dimension of the system or object (i.e., mass) under observation. Generally speaking, there are two principal categories of length: the macroscopic and the microscopic. The prefixes "micro-" and "macro-" come from Greek words that mean "small" and "large," respectively. The suffix "-scopic" originates with the word "scope," which in turn originates with the Latin word "scopus," which can mean aim, target or object of attention.

> **NOTE:** *The terms 'microscopic energy' and 'macroscopic energy' could be replaced here with 'microscopic viewpoint or approach' and 'macroscopic viewpoint or approach'.*

The macroscopic context considers a certain quantity of matter without considering the events occurring at the molecular level. The macroscopic approach is based on the study of the overall behaviour or gross behaviour of a number of molecules. Hence, the macroscopic "forms" of energy are those a system (or object) possesses as a whole with respect to some outside reference frame (e.g., the earth). The macroscopic viewpoint maintains:

1. The energy expressed/possessed by the gross or average behavior of multiple molecules that can be explained based on the continuum assumption.
2. Time average influence of multiple molecules.
3. Effects can be perceived by senses and/or measured by instruments, such as the effects of pressure and temperature.
4. Substances perceived as infinitely divisible or existing along a continuum.

The microscopic "forms" of energy, on the other hand, are those related to the molecular structure (architecture) of a system, its molecular bonding, and the degree of the molecular activity, which are independent of outside reference frames. The microscopic context maintains:

1. The energy expressed/possessed atoms or molecules are considered based upon statistical considerations and probability theory in connection with a model of the atom.
2. The 'internal energy' of a system is the sum of its microscopic forms of energy.

The macroscopic and microscopic energy value of a system can be combined to express the 'total energy' of a system - the sum of its macroscopic energy and microscopic energy.

Kinetic energy and the spatial length of a system in question can be combined in value. Kinetic energy is the energy of motion: the motion of large objects (macroscopic kinetic energy), or the movement of small atoms and molecules (microscopic kinetic energy). Macroscopic kinetic energy is sometimes referred to as "high quality" energy, while microscopic kinetic energy is more disordered and "low-quality". Macroscopic kinetic energy can become microscopic kinetic energy through 'friction'.

19.3.3 Classified by: Spatial medium

It is customary to say that energy exists in different forms which are transformed or converted into one another during physical processes. However, using the term "energy form" for the respective categories is unsatisfactory because it easily leads to the misinterpretation that there are different kinds of energy. The term 'energy carrier' more accurately accounts for the well-known but little recognized natural law that energy always flows simultaneously with at least one other existent (physical) quantity. Hence, the term 'energy carrier' is able to provide clear language of how energy is transported, exchanged, and stored. The substance-like physical quantity which flows while energy is flowing, "carries" the energy. It is imprecise to speak about the forms of something that itself does not change, but rather, which only changes carriers.

Of course, there are limits as to how literally the expression "energy carrier" should be understood. The word "carry" implies here only a temporal relationship between the flow of energy and the flow of an energy carrier. It is not meant to imply that energy and its carrier necessarily occupy the same position in space or even flow with the same velocity. An energy carrier can be "loaded" with more or less energy in the same sense that a carrier of commodities, say a pickup truck, can be loaded with more or less of a commodity.

The picture of energy carriers and energy load factors is especially useful to describe devices which are traditionally called "energy transformers" or "converters." Traditionally speaking, energy flows into an energy transformer in one form and out in another. Unfortunately, this way of speaking suggests that one physical quantity is transformed into another within such a device. Actually, however, the energy simply changes its carrier within the device. In other words, the energy is transferred from one carrier to another within the device. Accordingly, the name energy transceiver is more appropriate to the actual function of such a device.

It is easy to graphically represent the energy transport from one device or region of space to another with the help of an energy flow diagram. Such diagrams provide the means for a simple, graphical calculus applicable to the solution of energy-related problems. A

19.3.4 Energy carriers, mediums, and forms

CLARIFICATION: *Energy can be absorbed by carriers and transferred between carriers.*

Energy can be classified according to the spatial medium (form or source) for which the quantity is being calculated. The following are not different forms of energy -- in physics, there is just energy, not different forms of energy. In particular, in physics, the conservation law states that there are no sources and there are no sinks of energy. There is such a thing as 'dissipation', which is covered by entropy per a thermodynamic system.

Energy may be classified for the following systems:

1. **Mechanical energy** - energy possessed by the mass of an object due of its motion and/or potential to move. Energy in the mass of an object due to its motion and gravitational position - the sum of the kinetic energy and potential energy of a mass.
   A. Mechanical kinetic energy - energy due to motion (moving objects)
   B. Mechanical potential energy - energy due to tension of objects.
2. **Chemical energy** - energy possessed by the composition of a chemical substance due to the condition and/or potential condition of its atoms. Energy in the chemical bonds of atoms and molecules. Chemical energy is released out of the reaction taking place between elements/molecules to form a more stable compound.
   A. Chemical potential energy - energy due to molecular position in a chemically-bonded "force" field (i.e., due to chemical bonds).
3. **Nuclear energy** - energy possessed by the composition of an atom due to the condition and/or potential condition of its binding nucleons. Energy in the nuclear bonds of an atom. Nuclear energy is the energy released by either a fission or a fusion process. Here again the resulting products

assume a more stable condition.

- Nuclear potential energy - energy due to dielectric position in an atomically-bonded “force” field (i.e., due to nucleus of atoms).

4. **Electric[al] energy** - energy possessed by the electric composition of an atom due to the condition and/or potential condition of its polar charge. Energy expressed in/as charged motion is electrical energy.
   A. Electric[al] potential energy - energy due to atomic position in an electrically-bonded “force” field.
   B. Electric[al] kinetic energy - when matter that carries a charge moves in an electric field it carries kinetic energy.
5. **Electromagnetic energy** - energy possessed by the composition of ether (electric and magnetic fields) due to the condition and/or potential condition of its dielectric inertial plane. Electromagnetic energy is a dynamic form of energy that is caused by the acceleration or oscillation of a charged particle. All substances above absolute zero (0 Kelvin) emit a range of electromagnetic energy.
   - The acceleration/oscillation of electrical charges produces dielectric radiation.
6. **Gravitational energy** - energy possessed by the mass of an object due to its motion and/or potential to move toward a zero inertial plane.
   A. Gravitational potential energy - energy due to mass position in a gravitationally-bonded force” field (i.e., due to height and weight in gravity field).
7. **Thermal energy** - the total energy of motion, rotation, and vibration of the atoms and molecules inside an object. Thermal energy is not the energy of a whole object itself (mass) moving. In a gas or gas mixture, like air, the motion (and rotation) of individual gas particles makes up this energy. In a solid, like a table, the thermal energy exists as vibration of atoms or molecules. Note that total thermal energy also includes some atomic forms of potential energy. The temperature of an object is determined by its total microscopic kinetic energy. Thermal energy is the result of molecular agitation causing rise in temperature. It is the basic energy form, in the sense all other forms of energy can be completely converted into thermal energy. The other way, i.e., the complete conversion of thermal into other forms of energy is not possible and is governed by the Second law of thermodynamics. Thermal energy is the energy of a vibrating molecule in all degrees of freedom (translational, vibrational, rotational, potential).

19.3.5 Movement and Oscillation of energy carriers

The oscillation and/or acceleration of an energy carrier will produce a related type of energy transfer:

1. Atomic and molecular oscillation generates thermal radiative electromagnetic energy (heat).
2. Mechanical oscillation generates acoustic/cymatic radiative energy (sound).
   A. Linear acceleration of matter results in force (thrust or propulsion).
   B. Circular acceleration of matter results in torque.
   C. Divergent acceleration of matter results in pressurization or explosion.
   D. Convergent acceleration of mater results in depressurization or implosion.
3. Electric [charge] oscillation generates electromagnetic radiative energy (light).
   A. Moving charges have a magnetic field.
   B. Linear acceleration of electric charge generates a pulse of electromagnetic radiative energy.
   C. Circular acceleration of electric charge generates electromagnetism.
4. Magnetic oscillation [in the presence of a conductor] generates electrical radiative energy [around the conductor].
5. Changing magnetic fields generate electric currents in a conductor.

NOTE: *All digital communications and wireless charging, at the present, are based on electron oscillation. Therein, the term ‘signal’ is another word for the propagation of energy conveying communicative information (i.e., a communication)*.

19.3.6 Sub-classified by: Pressure gradient

NOTE: *There are pushing pressures and there are pulling pressures.*

Every aspect of power production and/or energy transfer in the world is about moving pressures from one place to another. Therein, flow exists because of a difference in pressure -- a difference in pressure between one point\location [in a conduit/medium] and another point\location [in the conduit/medium]. When there is no pressure [gradient], there is no flow of energy, and hence, no power.

INSIGHT: *Without pressure, there is no movement. Without cause, there is no pressure. Without intention, there is no cause. Without movement, there is no intention. Without desire, there is no suffering.*

When a pressure gradient has been established, energy always flows spontaneously, and without the need for work, from areas of high to low pressure (given a conducive medium for flow).

Pressure gradients:

1. Mechanical pressure gradient (strain)
2. Molecular pressure gradient (temperature, thermal)
3. Atomic pressure gradient
4. Electrical pressure gradient (voltage)
5. Magnetic pressure gradient (gauss)
6. Electromagnetic pressure gradient (electron-volt)
7. Electrochemical pressure gradient

Pressure types:

1. Electrical pressure (voltage) = energy / unit charge
2. Mechanical pressure (stress) = force / unit area
   A. Acting parallel (shear stress, the force is termed 'shear force')
   B. Acting perpendicular (normal stress\fluid pressure)
   C. The degree of mechanical deformation is termed 'strain'
3. Fluid pressure = energy / unit volume

Absence of pressure:

1. No hydro pressure = no flow of water = no hydro current = no hydro power
2. No electric and/or magnetic pressure = no flow of charge = no electric current = no electricity or electrical power
3. No mechanical pressure = no force or torque = no change of position = no mechanical power
4. No molecular pressure = no temperature = no thermal current = no thermal power

Water analogy describing the same principle, but different processes:

1. Matter flowing as water molecules (water molecules are flowing) - There are "excess" water molecules on one end of a conduit that are able to flow to the other end that has less molecules.
2. Matter flowing as charged particles (electron or proton particles are flowing) - There are "excess" charged particles (electrons) on one end that are able to flow to the other end that has less charged particles (electrons).
3. Ether flowing as dielectric fields - There are "excess" fields on one end that are able to flow to the other end that has less dielectric fields.

19.4 Energy Transfer modes

> **CLARIFICATION:** *It is not "the energy" being transported through the electromagnetic field, a fuel line, or a house wall which has different characteristics, but rather, a substance-like physical quantity which flows simultaneously with the energy in each case. Consequently, energy is not actually transformed or converted within a so-called "energy transformer" or "energy converter." Rather, it is correct to say that the other substance-like physical quantity that flows along with the energy is exchanged/transferred/converted within such a device. For example, energy is brought into a power plant together with coal and oxygen or, scientifically speaking, together with the amount of substance (the quantity measured in moles) of coal and of oxygen, and energy always flows out of the power plant simultaneously with electric charge.*

Energy transfer refers to the movement or flow of energy from one area of space-time to another without changing its form. There are different processes that allow for the transfer of energy; these different processes are otherwise referred to as 'modes of energy transfer' (or 'energy transfer modes'). These transfer modes (processes) are generally signified linguistically as gerunds (verbs functioning as nouns). The presence of the process (or mode) means the transfer of energy. Recall that the time rate of energy transfer is called 'power'.

> **NOTE:** *Although every transfer will mean some conversion to heat, the transfer of energy may or may not mean the complete conversion of the medium participating in the transfer. Whereas a combustion system involves some conversion of the input medium, a wind turbine system does not involve conversion of the wind to some other medium. Energy can be transferred without changing carrier, such as when [electrical] energy "moves" along a wire, or [thermal] energy "moves" from the inside of a hot cup to the outside.*

The mechanisms of energy transfer at a system boundary are:

1. Temperature (heat or heating) - heat transfer
2. Pressure (work or working) - work transfer
3. Volume (mass flow or permeating) - mass transfer

> **NOTE:** *In a closed systems (i.e., systems with a fixed mass), energy can only transfer across the boundary as heat and/or work, not mass flow. The boundary of an open systems allows for energy transfer by mass flow, as well as temperature (heat) and pressure (work).*

Systems may be said to "possess" energy, but they cannot be said to "possess" heat or work. These are modes of transfer, otherwise known as 'boundary phenomena', because they are recognized at the boundary between the two or more systems. The transfer modes are associated with a [boundary] process, not a state of the system. Also, they are path dependent functions (i.e., their magnitudes depend on the path followed during a process as well as the end states).

Heat and work are energy in transit. Thus, heat and work are not properties of state, but energy that is in transport across system boundaries, to or from the environment. It is not possible to measure how much heat or work are present in an object, but rather only how much energy is transferred among objects in certain ways during the occurrence of a given process. Work and heat are, in a way, transport concepts for energy.

Heat and work are measured as positive (additive) or negative (subtractive) depending on which side of the transfer they are viewed.

These transfer processes can be viewed from several perspectives. From the perspective of gradient flow, there are two primary transfer modes:

1. Differences in temperature lead to the transfer of energy.
2. Differences in pressure lead to the transfer of energy.

The are a fixed number of ways that energy can be transferred (i.e., “exchanged”). These transfer processes can be viewed from the motion of a medium:

1. **Displacing matter** - by accelerating matter through a pressure differential.
2. **Permeating matter** - by displacing the location of some matter into other matter through pressure differential.
3. **Pressurizing matter** - by compressing matter through pressure differential.
4. **Heating matter** - by [rapidly] accelerating the atoms and molecules of matter through pressure differential.
5. **Waving matter (mass radiation)** - by colliding matter, creating compression and rarefaction of mass (i.e., “waves”). By accelerating an interface to a closed and non-displaced system.
6. **Waving ether (electromagnetic radiation) -** by accelerating charges, creating compression and rarefaction of ether (i.e., “waves”).
   NOTE: *Waving is transfer of energy without the transfer of matter. A wave is a vibration (vibratory pressure as compression and rarefaction) that transfers energy from one place to another without transferring matter (solid, liquid or gas).*

Energy can be transferred into our out of a system in several ways:

1. Through mechanical interactions - energy transferred mechanically as work (i.e., in a coherent manner)
2. Through thermal interactions - energy transferred thermally as heat (i.e., in an incoherent manner)
3. Through radiation interactions - energy transferred

NOTE: *Any of these transfers increase the system’s internal energy (per ΔE).*

Energy can be transferred into our out of a system in several ways:

1. **Working - work transfer mode**: The coherent and directed transfer of energy by a force causing a displacement at the point of application of the force. Work transfer (working) may be contrasted with heat transfer (heating). Work transfer refers to all transfers that do not involve a temperature difference. May be viewed as the flow of “non-thermal energy transfer”. If work is *done on* a system, then the energy of the system increases (input energy transfer). If work is *done by* a system, then the energy of the system decreases (output energy transfer).
2. **Heating - Heat transfer mode:** The transfer of energy driven by a temperature difference between two regions in space. Heat transfer takes place due to the presence of a difference in temperature. Heat always flows from the system with a higher temperature to the system with a lower temperature. Heat is the quantity of energy which crosses the boundaries of a thermodynamic system. When added to a system heat transfer causes the energy of a system to increase, and heat transfer from a system causes the energy to decrease.

There are several important principles to note when discussing these two transfer modes:

1. Work transfer can be completely converted into heat transfer.
2. Heat transfer cannot be completely converted into work transfer, because some of the “energy” will disperse irrecoverably. The degree to which the “energy” is dispersed is known as the entropy of the system.
3. Work transfer will always produce some heat transfer.
4. Both heat and work are path functions, and vary with the manner in which the process is carried out.
5. A system cannot contain or store either heat or work.
6. Heat into a system and work out of a system are considered positive quantities.

Energy can be transferred into our out of a system in several ways (modes) - it must flow or it is not energy transfer:

1. **Mechanical radiative transfer (wave flow)** - energy is "transferred" by the propagated disturbance of a physical medium (flow of "mechanical energy" as wave collisions). The flow of a propagating [contraction and rarefaction] disturbance of space. Matter does not flow with the transfer of energy.
2. **Matter transfer (mass flow)** - energy is transferred by the movement of physical matter through the boundary of a system, carrying energy with it. The flow of matter as volume transfer. As mass flows into a system, the energy of the system increases by the amount of energy carried with the mass into the system. Mass leaving the system carries energy with it, and the energy of the system decreases. Since no mass transfer occurs at the boundary of a closed system, energy transfer by mass is zero for closed systems.
3. **Electromagnetic transfer (field flow)** - energy is transferred by the propagation of an electromagnetic field (Read: electromagnetic induction or radiation). The flow of "electromagnetic energy". Or, the flow of ether as a propagating [contraction and rarefaction] disturbance of counterspace. The transfer processes are known as:
   A. Electromagnetic induction.
   B. Electromagnetic radiation.
4. **Electrical transfer (charge flow)** - energy is transferred by electric charge/current propagating through conductive matter. The flow of "electrical energy". The transfer process is known as:
   A. Electrical [charge] conduction.

19.4.1 Transfer types

Energy can be transferred from one point in space-time to another in three ways:

1. Through the action of forces.
   B. **Electric and magnetic force fields** - Charged particles, upon which electrical fields exert forces, possess potential energy in the presence of an electric field in a way similar to that of an object in a gravitational field. These force fields can accelerate particles, converting a particle's potential energy into kinetic energy. Likewise, charged particles can interact via the electric and magnetic fields they create, transferring energy between them, and in the case of an electrical current in a conductor, cause molecules to vibrate, i.e. converting electrical potential energy into heat.
   C. **Frictional Forces** - The macroscopic (large-scale) energy of an object, that is, the potential and kinetic energy associated with the position, orientation, or motion of the entire object, not counting the thermal or heat energy of the system, can be converted into thermal energy (heat), whenever the object slides against another object. The sliding causes the molecules on the surfaces of contact to interact via electromagnetic fields with one another and start vibrating.
   D. **Gravitational force (possibly, magnetic)** - When gravity accelerates a falling object it converts its potential energy to kinetic energy. Likewise, when an object is lifted, the object stores the energy exerted by the lifter as a potential energy in the earth-object system.
2. When atoms absorb or emit electromagnetic radiation (i.e., photons of light). When light falls on an object, an incident photon may either pass through the object, be reflected by the object, or be absorbed by the atoms making up the object. If most of the photons pass through, the object is said to be transparent. Depending on the smoothness of the surface on the scale of the photon's wavelength, the reflection may be either diffuse (rough surface) or coherent (smooth surface). If the photon is absorbed, the photon's energy may also be split up and converted in the following ways:
   A. **Photothermal effect:** the energy absorbed may simply produce thermal energy, or heat in the object. In this case the photon's energy is converted into vibrations of the molecules called phonons, which is actually heat energy.
   B. **Photoelectric effect:** the energy absorbed may be converted into the kinetic energy of conduction electrons, and hence electrical energy.
   C. **Photochemical effect:** the energy may bring about chemical changes which effectively store the energy.
3. When nuclear reactions occur, that is, when there are rearrangements of the subatomic particles that make up the nuclei of atoms. There are two basic types: **Fission** - when nuclei combine, and **Fusion** - when nuclei split apart.

19.4.2 Transfer (carrier) interactions

Electromagnetic waves (light) and mechanical waves (sound) interact with physical materials in various ways that impact their "transfer" of energy. There are three principal ways in which waves (compression and rarefaction of space or counterspace) interact with matter; these are known as wave behavior interactions:

1. **Transmission (light wave transmission or mechanical wave transmission):** The passing of a wave through a material. For instance, light passes through an object - an object is either transparent (the light passes straight through), or translucent (the light passes through, but its direction "scattered" by the material). Light rays that pass through an interface are transmitted rays. These rays bend. This bending is called refraction. The direction and magnitude of refraction depends on the relative densities of the two media and the angle of incidence.
2. **Reflection (light wave reflection or mechanical wave reflection):** The bouncing back of a wave after it strikes a barrier. For instance, light bounces off a surface. Reflection can either be coherent (the angle of incidence equals the angle of reflection) or diffuse (the reflected direction is scattered).
3. **Absorption (light wave absorption or mechanical wave absorption):** The transfer of energy from a wave to matter as a wave passes through it. For instance, light enters a material and does not pass through. Instead, its energy is converted into "thermal energy" as microscopic vibrations of the material, or is absorbed by chemical reactions triggered by the light (the photochemical effect).

Consider a passive solar home in the winter with sunlight propagating to a window on the home. The following may be said about the system:

1. Light is transmitted through the window (which is either translucent or transparent) .
2. The light contacts the floor and is either reflected or absorbed (after several reflections around the room, almost all is eventually absorbed. A tiny bit is reflected back out the window.
3. The floor (and other surfaces where the light contacts), are "heated" by the absorption of light.
4. Some of this heat conducts into the material.
5. Some of this heat is re-radiated (at infrared wavelengths), back into the room.
6. Air near the surfaces is "heated" by this re-radiation and by contact (conduction) with the wall.
7. The "heated" air rises (convection).

19.4.3 Work transfer mode

IMPORTANT: *Modern physics is still confused over the concepts here as is visible in the following: Work and energy are in the same units (joules), but work is always a change in energy.*

Work describes a transfer change in energy; it's not a conserved quantity in itself unless embedded in the concept of energy, which is conserved. With the above said, the "ability/capacity of a system to perform/do work" is a common and imprecise definition of energy. As a measure, energy is the total amount of work that can be done.

Work is the difference in energy between when you started, and when you're done. "Work" is just the amount of energy added (the amount of energy it has at the end, minus the amount of energy it had when you started).

In physics, it is said that whenever a force is applied that causes motion, then "work" is done. Work is a mode of transferring energy; it is an energy interaction between a system and its surroundings. Work is the use of force to act on an object in order to move that object in the same direction as the force. As a principle, work has two expressed attributes. Firstly, work is a transfer of directed/ordered energy, as contrasted with "disordered" energy (i.e., heat). Work refers to an activity involving a force and movement in the direction of the force.

Work is done on something or by something. Another way of understanding work is that work is a change in energy via a force. There are other ways of changing energy, such a through thermal contact. The point is that work changes the energy of something and the change is through application of a force (or torque).

Hence, if there is force, but no movement of distance/displacement, then there is no work. Here, a '**force**' is an action that pushes or pulls (applies a pressure to) an object or material substance or physical system.

- Work (J) = force (N) x distance (m)
- 1 Joule = 1 newton x 1 meter
- In units energy = joules (J)
- Work = joules (J)

Work is scalar. Work is not a vector, but force and displacement are vectors. Work is the product of a component of a force on an object times the displacement of the object, while the force is being exerted in the direction of displacement.

- $W = {}^{+}F \times {}^{+}d$ - when force causes a displacement, work (energy) is positive (F x d = work)
- $W = {}^{+}F \times {}^{-}d$ - when force hinders a displacement, work (energy) is negative ($F \times {}^{-}d = {}^{-}$work).
- $W = {}^{+}F \times {}^{0}d$ - when force results in no displacement, there is no work (F x 0 = 0 work).

Secondly, work (W) is accomplished by a force (f) acting through a distance (d).

- $W = \int f_i \cdot \Delta x_i$ (i.e., Work = Force x Distance)
- $W = F \times \Delta x$
- where, x is displacement

For a constant force; the work done formula as force x distance, only applies if you have a constant force:

$W = f_i \cdot \Delta x_i$

From a systems perspective, work is the amount of directed energy transferred into or out of a system. Work can increase or decrease energy. In most classical scenarios, work will change either the kinetic or potential energy. In concern to work as a factor of potential or kinetic energy:

- $W = \Delta K$ (kinetic)
- $W = -\Delta U$ (potential)

19.4.4 Heat transfer mode

IMPORTANT: *Modern physics is still confused over the concepts herein as is visible in the circularity of their definitions: (1) Energy does work or produces heat; (2) but, heat is a transfer of energy.*

Heat, like work, is neither a thing, nor a form of energy. Heat is not a substance that is being transferred [between any two things]. Heat is the energy in transit from one body to another under the influence of a temperature gradient. Heat cannot be stored as such. Heat flowing into a body merely changes the internal energy of the body. The internal energy of the carrier is increased after heating. The temperature of a body gives a measure of the intensity of heat.

NOTE: *The enthalpy of a system is the sum of the internal energy and the pressure-volume of the system. It is a function of the state of the system and depends on the temperature and pressure of the system. The absolute value of the enthalpy of a substance cannot be calculated, but values relative to some arbitrary chosen reference state can be determined.*

When a temperature difference exists across a boundary, the second law of thermodynamics indicates the natural flow of energy is from the higher temperature body to the lower temperature body.

DEFINITION: *The* **specific heat capacity** *of a substance is defined as the amount of heat required to increase its temperature by one degree. Specific heat capacity is measured in joules per kilogram degree-celsius.*

Before it is transferred, the energy which remains within the boundary of the system is not heat, but 'internal energy' or 'total available energy'. Once a system absorbs heat, the latter is no longer heat, but internal energy of the system. In other words, it stops being heat because heat is no longer being transferred between two systems at different temperatures. For it to be heat it must be in the process of being transferred from one system to another system. Heat cannot be stored nor contained by any system because heat is a process function.

NOTE: *An adiabatic process is one in which the system is perfectly insulated and the heat transfer is zero.*

A process function, or process quantity, is a physical quantity that describes the evolution or shift through which a thermodynamic system passes from an initial equilibrium state to another equilibrium state. It is a category error to use the expression "heat stored" if one does not clarify that it is not heat (a process) which is being stored, but energy which has been transferred from one system to another transposing the boundary of the acceptor system. The proper expression for this is "energy stored by heat transfer"; or simply "energy stored".

The energy emitted or released by the system becomes heat the moment it crosses the boundary of the system (i.e., the moment it becomes energy in transit). Remember, process quantities cannot be stored or contained because they describe the trajectory by which a system acquired an equilibrium state. A process function or process quantity is not a state function. A state function is a property of a thermodynamic system which depends only on the current state of the system. Internal energy is a state function.

There is no such identifiable thing as "heat energy" in an object. Hence, "hot" is not a substance. Instead, heat is a mode of energy transfer representing the flow/transfer of energy spontaneously (across systems or within a system) due to temperature differences. When a suitable physical pathway exists, energy flows spontaneously from a hotter to a less hot (i.e., "colder") body. The name of the transfer process is heat transfer. What gets transferred is a quantity of energy. Heat describes what energy is doing at a given time. Essentially, heat is any energy transfer that is not macroscopically ordered (i.e., an energy transfer expressing all degrees of freedom; it "disperses"). When energy disperses, it is not destroyed, but rather, that it is lost for useful purposes.

19.4.5 Temperature (direction of heat transfer)

Temperature is a property of matter that determines the direction heat will flow when two mediums are brought into contact. The direction of heat transfer is based on temperature. In order to have a change in temperature there must be a transfer of energy between systems.

Temperature is a property which is directly proportional to the kinetic energy of the substance under examination.

Temperature is measured in either:

1. Kelvins (K) with zero motion as its reference point (0°K).
2. Celsius (C) with the freezing point of water as its

reference at (0°C).

The Celsius scale is a derived scale, defined in relation to the Kelvin temperature scale. The Celsius scale is an interval system, not a ratio system; it follows a relative scale and not an absolute scale. This can be seen because the temperature interval between 20 °C and 30 °C is the same as between 30 °C and 40 °C, but 40 °C does not have twice the air heat energy of 20 °C. A relative scale adds an unnecessary degree of abstraction (generating the potential for confusion) over an absolute scale. An absolute scale represents as close [an expressed] alignment with reality as is possible given what is presently known.

Heat is energy in transit, it is dynamic in nature and heat flow stops only at their equilibrium temperature state (i.e., thermal equilibrium).When two bodies are in thermal equilibrium with a third body, then they must be in thermal equilibrium with each other. This is called the Zeroth Law of Thermodynamics and is the basis for temperature measurements, since the thermometer must come to thermal equilibrium with the object being measured.

19.4.6 Modes of heat transfer

Heat transfer is energy in transition across the system boundary due to a temperature difference, there are three modes of heat transfer at the boundary that depend on the temperature difference between the boundary surface and the surroundings. These are:

- Conduction
- Convection
- Radiation

There are three types of heat transfer (i.e., three mechanisms by which energy is transferred via heat). The first two (conduction and convection) refer to the direct transfer of energy, whereas the radiation is a conversion of energy to a different form (electromagnetic radiation, light), and the subsequent travel (transfer/transport) of that radiation. These heat transfer processes may also be referred to as modes:

1. **Conduction (conductive heating/thermal conduction)** - transfer of thermal energy through an object/substance by atomic movement. Direct transfer by contact through a solid or stationary fluid. When a temperature gradient exists in a stationary medium, which may be a solid, liquid, or gas, the term "conduction" refers to the heat transfer that will occur across the medium.
   A. A **thermal/heat conductor** is a substance/material that allows thermal energy to move through itself easily. Materials that do not allow thermal energy to move through them easily are called **thermal/heat insulators**. Similarly, electrical conductors allow electrical energy to move through easily, while electrical insulators do not allow electrical energy to move through easily.
2. **Convection (convective heating/thermal convection)** - transfer of thermal energy from a surface to a moving fluid by movement of groups of molecules. Convection refers to the heat transfer that occurs between a surface and a moving fluid when they are at different temperatures. Simplistically, convection is the movement of a fluid in response to heat. A substance experiencing convection will move in the form of a [convection] current. Convection takes place through advection, diffusion, or both.
   A. **Advection** is the movement of some material dissolved or suspended in the fluid. For instance, if pure water is heated, there will occur convection of the water. Advection cannot occur because there is nothing dissolved or suspended in the fluid to advect. If silt is suspended in the water, and the liquid mixture is heated, then there will occur convection of the water and advection of the silt.
   B. **Diffusion** is the net movement of particles from high concentration to low concentration.
3. **Radiation (radiative heating/radiant heat/ thermal radiation)** - transfer of thermal energy through space-time from the emission and absorbance of electromagnetic waves - the emission of electromagnetic radiation and its absorption. All objects, even those that are in equilibrium (at equal temperature) with their surroundings, continuously emit/radiate electromagnetic waves (i.e., "light waves") into their surroundings. The source of this radiation is the thermal energy of the materials, the movement of the object's molecules. Radiation does not require matter, unlike convection and conduction.

19.4.7 Electromagnetic transfer mode

Electromagnetic transfer mode is the transfer of energy as electromagnetic fields at near the speed(s) of light. Electromagnetic energy can be reflected or emitted from objects through electric and/or magnetic waves traveling through space. The electromagnetic spectrum is the range of all types of electromagnetic radiation (electromagnetic radiation being a type of electromagnetic energy). Electromagnetic radiation is a kind of energy that travels and spreads outward as it travels. The electromagnetic spectrum is a categorization of all electromagnetic waves by frequency, wavelength or photon energy. Electromagnetic frequencies are produced where electricity flows. One does not exist without the other. Electricity, understood as the movement of electrical charges, generates an EM-waves, relative to the geometric conditions of the circuit and frequency conditions in the current flow. Conversely, an electromagnetic wave can generate electricity (photoelectric effect). Electromagnetic radiation can be described in terms of its wavelength — the distance between the crests of the waves — or its frequency — the number of crests that pass by a fixed point during a fixed time interval.

> **DEFINITION: Coupling** *is the transfer of energy from one medium, such as a metallic wire or an optical fiber, to another medium (which may be of the same composition, but separated by space). In an electrical circuit, coupling is the transfer of electrical energy from one circuit segment to another.*

Electromagnetic transfer mode refers to the transfer of energy between two spatially separated objects -- the energy is transferred without contact. Electromagnetic "wireless" power (wireless energy transfer) techniques fall into two categories:

1. Radiative (radiation-based)
2. Non-radiative (induction-based)

Therein, there are three kinds of wireless power transfer technology (wireless transmission of electricity and/or "energy") in accordance with its working principles (technically, these are all forms of electromagnetic transfer):

1. **Electromagnetic and electrodynamic induction (non-radiative)**
   - **Magnetic induction** - magnetic field coupling mode
2. **Electrostatic induction (non-radiative)**
   - **Electric induction** - electric field coupling mode
3. **Electromagnetic radiation (radiative)** - electromagnetic field coupling mode

> *Take note that each mode/method listed above has multiple different names in the literature and in application.*

19.4.8 Magnetic induction mode

> *A.k.a., Magnetic field coupling mode; direct induction; electromagnetic induction*

This is a method of producing electromotive force (voltage) and/or heat across an electrical conductor due to its dynamic interaction with a magnetic field. The magnetic field may come from either moving permanent magnets or alternating current electromagnets.

Electromotive force causes the movement of electric charge. How much electromotive force is present between two points in a circuit is measured in units of 'volts'. The electrical current that flows in this situation is known as an 'induced current'. Electromagnetic induction occurs when a changing (moving/dynamic) magnetic field induces an electrical current in a closed loop. Note that any change in the magnetic field around a conductor will induce a voltage. The more voltage induced, the more electrical current produced (if an electrical circuit is present).

> **CLARIFICATION:** *Induction refers to energy transfer without contact (versus conduction, which is by contact). The basic process of generating electrical power with magnetic fields (and without contact) is known as* **induction***. This specific type of induced current process is also called magnetic induction to distinguish it from charging by induction, which utilizes the Coulomb force. Inductive charging is also known as wireless charging.*

Electromagnetic induction relies on 'magnetic flux'. Magnetic flux refers to how a conducting material (or any material) is affected by a magnetic field. Magnetic flux is the product/strength of the magnetic field multiplied by the conductive surface area perpendicular to that magnetic field. Electromagnetic induction occurs when there is a change (change only) in the magnetic flux (over time). By continuously varying the magnetic field or surface area (angle or volume) a continuous electromagnetic induction will occur. Wrapping/coiling wire is a good way to increase magnetic flux.

> **CLARIFICATION: Electromagnetic flux** *can be classified into 2 types: (1) Electric flux is defined as the number of field lines or the concentration of field lines of an electric field perpendicular to a surface. (2) Magnetic Flux is the number of magnetic field lines or the concentration of magnetic field lines perpendicular to a given surface.* **Electrodynamics** *is the branch of physics which deals with rapidly changing electric and magnetic fields. A current moving through a conductor creates both magnetic and electric fields. A time-varying current will produce time-varying fields. Time-varying currents are nothing more than a macroscopic series of charges undergoing time-varying acceleration.*

Magnetic induction is used in the following energy transfer applications:

19.4.8.1 AC electricity (in conductive coils)

AC electrical generators (alternators) - devices that use [electro]magnetic induction to produce electricity (as the flow of electric charge, electric power).

19.4.8.2 Heating (in ferrous metals)

Induction heating (e.g., induction cooking) - the process of heating an electrically conducting object (usually a metal) by electromagnetic induction (specifically, magnetic inductive coupling), through heat generated in the object by eddy currents (also called Foucault currents). Induction heating occurs due to electromagnetic force fields producing an electrical current in a part. By applying a high-frequency alternating current to an induction coil, a time-varying magnetic field is generated. The parts heat due to the resistance to the flow of this electric current. An induction heater consists of an electromagnet, and an electronic oscillator that passes a high-frequency alternating current (AC) through the electromagnet. The rapidly alternating magnetic field penetrates the object, generating electric currents inside the conductor called eddy currents. The eddy currents flowing through the resistance of the material heat it by Joule heating. Induction heaters are used to provide alternating electric current to an electric coil (the induction coil). The induction coil becomes the electrical (heat) source that induces an electrical current into the metal part to be heated (called the workpiece). No contact is required between the workpiece and the induction coil as the heat source, and the heat is restricted to localized areas or surface zones immediately adjacent to the coil. This is because the alternating current (AC) in an induction coil has an invisible force field (electromagnetic, or flux) around it. Furnaces (as an alternative method of heating) tend to be large, have long start-up and shut-down times, and emit fumes and by-products of combustion, both a pollutant and a potential safety hazard. The induction heater can be small and, as all electric devices, is immediately turned on and off. It is a “clean” process and safer for those operating the system. It also has fewer maintenance costs than furnaces. As with conduction heating, induction heating has the benefit that all of the power supplied goes directly into the workpiece and heating times are short. They fit well into automated production methods, are easily controlled, and the process is highly repeatable. There are some surprising benefits to induction heating. For example, alloys are easily mixed in induction heating processes because the induced field automatically stirs the melted metal! Also, special techniques— precision melting, hardening of surface— can be implemented in the process. Induction heaters require electricity.

19.4.8.3 Voltage transformation (between electrically conductive circuits)

Electrical transformers (a.k.a., non-resonant inductive coupling; i.e., conventional transformer or electrical power distribution transformer) - devices that use electromagnetic induction to change the voltage of electric current. Electrical transformers transform one voltage into another voltage through electromagnetic induction. In other words, it is a device in which an input alternating current produces an output alternating current of different voltage. Note here that transformers work with AC, not DC. It is called an electrical transformer, because it transforms electrical energy into magnetic energy, then back into electrical energy again. A transformer’s main purpose is to transfer electrical energy from the primary coil to the secondary coil. A transformer’s basic operating principle: the transfer of power from the primary to the secondary circuit occurs via electromagnetic coupling. An electrical transformer is a form of wireless energy transfer. The primary and secondary circuit of ta transformer are not directly connected. Here, energy transfer takes place through a process known as mutual inductance (without any physical contact in between). In transformer theory, electromagnetic (EM) induction refers to the phenomena that electromagnetic changes in one place induce (EM) changes in another place.

19.4.9 Electrodynamic induction mode

> **Also known as:** *Magnetic field coupling mode (a.k.a., inductive coupling; magnetic coupling; magnetic inductive coupling; inductive power transfer; resonant magnetic induction; resonant inductive coupling*

This is the near field wireless transmission of [electrical] energy (by the transfer of electromagnetic energy) between two magnetically coupled coils that are part of resonant circuits tuned to resonate at the same frequency. The resonant inductive coupling process occurs in a resonant transformer, an electrical component which consists of two coils wound on the same core with capacitors connected across the windings to make two coupled LC circuits. Resonant transformers are widely used in radio circuits as bandpass filters, and in switching power supplies. Resonant inductive coupling is also being used in wireless power systems. Here the two LC circuits are in different devices; a transmitter coil in one device transmits electric power across an intervening space to a resonant receiver coil in another device. It is the transfer of energy between a current-carrying conductor and nearby conductors due to a time-varying magnetic field that is created by time-varying current in the energized conductor. Magnetic induction concerns electric currents generated by the motion of a magnetic flow along a conductor. In other words, magnetic field coupling is caused by the current flow in conductors. Magnetic field coupling is created by inductive means (inductive coupling). The magnetically induced current in each nearby conductor will be slightly different since it depends on the relative location of each individual conductor to the energized conductor. The coupling mechanism can be modelled by a transformer. According to the transfer distance, the magnetic field coupling mode can be mainly classified into short-range electromagnetic induction and mid-range strongly coupled magnetic resonance (SCMR). The transfer efficiency and transfer power of electromagnetic induction are normally high, but the transfer distance is limited to centimeter level. In contrast, the transfer efficiency and transfer power of SCMR are a marginally lower, but the transfer distance can achieve meter level to realize mid-range power transfer. Optimal for mid-range wireless power transfer. This technology is being developed for powering and charging portable devices such as cellphones and tablet computers at a distance, without being tethered to an outlet. Inductive power transfer works by creating an alternating magnetic field (i.e., an electromagnet; flux) in a transmitter coil and converting that flux into an electrical current in the receiver coil.

> **NOTE*:* Non-dynamic magnetic induction (induced magnetism)** *is the production of a magnetic field in a piece of unmagnetized iron or other ferromagnetic substance when a magnet is brought near it. The magnet causes the individual particles of the iron, which act like tiny magnets, to line up so that the sample as a whole becomes magnetized. Most of this induced magnetism is lost when the magnet causing it is taken away.*

19.4.10 Electrostatic induction mode

> **Also known as:** *Electric field coupling mode capacitive coupling; electrostatic influence*

This is the near field transfer of energy between an energized conductor and the nearby conductors due to a time-varying electric field that is created by moving charge in the energized conductor. It pertains to magnetic flows produced by an electric charge (voltage). It is caused by an electric field gradient (voltage difference) or differential capacitance between conductors. The phenomenon of producing induced charges is known as electrostatic induction. The principle itself refers to the redistribution of the surface charges on the object. In other words, it is the production of an unbalanced electric charge (i.e., static electricity) on an uncharged conductor as a result of a charged body being brought near it without touching it. In other words, it is a redistribution of electrical charge caused by the influence of nearby charges. If the charged body is positively charged, electric charge in the uncharged body will flow toward it; if the opposite end of the body is then grounded, electric charge will flow onto it to replace those drawn to the other end, the body thus acquiring a negative charge after the ground connection is broken. A similar procedure can be used to produce a positive charge on the uncharged body when a negatively charged body is brought near it. Electrostatic induction is an efficient way of using a charged object to give something a charge, of the opposite sign, without losing any of the original charge. Electric field coupling is capacitive in nature (capacitive coupling). Hence, the coupling mechanism can be modelled by a capacitor (a capacitor is defined as two conductors separated by a dielectric, and may be used to store charge, "electrical energy"). A high-frequency and high-voltage driver source excites the resonant transmitter to generate an alternating electric field which can couple with the resonant receiver. Energy will be delivered as soon as this coupling relation is set up. The transfer efficiency of this mode is affected by surrounding objects. Optimal for short-range wireless power transfer.

> **NOTE:** *Electromagnetic resonance uses "antennas", and electromagnetic induction uses "transformers".*

19.4.11 Electromagnetic radiation mode (EMR)

> **Also known as:** *Resonant coupling; electromagnetic resonance*

This is the far field transfer and receiving of electromagnetic energy. Electrical energy is generally converted into electromagnetic energy, which can be radiated outward (as "EM waves"), which are then received and converted back into electric energy with using a silicon rectifier antenna in the receiver. Electromagnetic radiation (EMR) is the emissive transmission of electromagnetic energy between two bodies not in contact (source to receiver). Electromagnetic radiation is used to transfer electrical energy (which may carry power and/or data) without an electrical conductor or inductive coupling.

> **NOTE:** *An antenna can be designed to react with either the electric or magnetic field of a an electromagnetic radiative wave.*

The transmitter and receiver are tuned to the same resonant frequency (to a mutual frequency). In general, this is accomplished through "radio waves" or optical laser devices. Electromagnetic radiation (i.e., "light") propagates by itself in a vacuum at very high speed (the speed/s of light). Because of its high power density and good orientation features, electromagnetic radiation mode is usually suitable for the long distance transfer applications. However, its transfer efficiency is severely affected by the material conditions (e.g., meteorological or topographical conditions), and the impacts on creatures and ecological environment are unpredictable. Optimal for long-range wireless power transfer.

> **NOTE:** *In electric circuits, this motivating force is voltage (a.k.a. electromotive force, or EMF). In magnetic circuits, this motivating force is magnetomotive force, or mmf. Magnetomotive force (mmf) and magnetic flux (Φ) are related to each other by a property of magnetic materials known as reluctance.*

19.4.12 Electrical transfer mode

Electrical conduction mode (a.k.a., electrical transfer mode; electrical conduction coupling, electron mode) is the transmission of electrical energy (as electricity - flow of electric charge carried by electrons conserved in a circuit) from a power source to an electrical load, such as an electrical power grid or a electrically powered device with the use of a conductor (physical contact). An electrical conductor is a substance in which electrical charges (e.g., electrons) move easily with the application of voltage (i.e., contains movable electrical charges). Electrical conduction can occur in a "wired" or "wireless" manner. Wireless power transfer (WPT; a.k.a., wireless energy transfer, wireless energy transmission, and wireless electrical transmission) is the transmission of electrical energy (as electricity) without the use of a discrete human-made (synthetic) conductor (e.g., atmospheric plasma channel coupling - air method; ground channel coupling - ground method. Electrical power transfer (EPT; wired power transfer) is the transmission of electrical energy (as electricity) with the use of a discrete human-made (synthetic) conductor (e.g., hard-wire using a wire; resistive using a resistor).

Energy can be transferred by electrical transmission. Within a wire this is accomplished through electric fields associated with electrons in the metal wire. The electrons literally push on each other, and convey force through the wire, which thereby transfers energy. For example, the electro-chemical processes in a battery create positive and negative electric charges at the battery contacts which push on, and hence force, the movement of electric charge. Electrical energy is converted to heat when some of the electrons encounter resistance - that is, when the electrons are pushed through materials causing heat, that is, cause the atoms of the material to start vibrating. Alternatively, the movement of electrons may give rise to electric and magnetic fields (such as in coils of a motor), which do work, such as turning the motor shaft.

> **NOTE:** *Bearings provide a convenient support for rotating shafts.*

Electrical energy can be transmitted by means of electrical currents made to flow through naturally existing conductors, specifically the earth, lakes and oceans, and through the upper atmosphere starting at approximately 35,000 feet (11,000 m) elevation — a natural medium that can be made conducting if the breakdown voltage is exceeded and the constituent gas becomes ionized. For example, when a high voltage is applied across a neon tube the gas becomes ionized and a current passes between the two internal electrodes.

> **NOTE:** *Whenever an electric current flows through a conductor, a magnetic field is immediately brought into existence in the space surrounding the conductor. It can be said that when electric charges are in motion, they produce a magnetic field. The converse is also true (as in, when a magnetic field embracing a conductor moves relative to the conductor, it produces a flow of electric charge in the conductor).*

20 Units and formulas for: Energy

In concern to watts, a quantity of energy is measured in watt-time (e.g., watt-seconds, or more commonly, watt-hours). Watt-hours means watts multiplied by the hours the watts are transferring energy (i.e., doing work) to form a total quantity of energy transferred, or potentially transferred. Similarly, watt-seconds means watts multiplied by the number of seconds the watts are transferring energy. Watt-hours are a measurement of energy, describing the total amount of energy (electrical, mechanical, etc.) used over time. Watt-hours are a combination of how fast the energy is used (watts) and the length of time it is used (hours).

> **QUESTION:** *How much energy does a system require to operate for one second, and one hour? The answer will come in units of watt-seconds and watt-hours.*

Watt-time (e.g., watt-hours) is a quantity of energy -- the quantity of energy transferred (or, work done) in a given amount of time. For example, watt-hours is a combination of how fast the power (e.g., electricity) is used (watts) and the length of time it is used (hours):

- Watt-time = Watts x time
- Wt = W • t

In hours:

- Watt-hour = Watts x Hours
- 1 Watt-hour = 1 Watt x 1 hr
- 1Wh = 1W • 1hr
- Wh = W • hr

> **NOTE:** *The unit for watt-hour may be abbreviated: Wh, W.h, or W-h (or in seconds, Ws, W.s, or W-s).*

More commonly, energy is measured in kilowatt-hours, the equivalent of 1000 watts of power for 1 hour.

- Kilowatt-hour = Kilowatts x Hours
- 1 Kilowatt-hour = 1 Kilowatt x 1 hour
- 1kWh = 1kW • 1hr
- kWh = kW • hr

In concern to the quantity of energy used by a load, for instance, a 60-watt (power) light bulb running for 1 hour (time), will have used 60 watt-hours of energy.

1. A 60W light bulb - requires 60W of power to run for 1 second.
2. Running for 1 hour.
3. Will use [a quantity of] 60 watt-hours of energy:
    - 60W x 1hr = 60Wh = .060kWh
4. 1 second is 1/3600 of an hour; hence, in 1 second a 60W light bulb uses:

- .060kWh x 1/3600s = .00001666kWh

In other words, a light bulb with a power rating of 60 watts will use 60 watt-hours per hour, or 60 watt-seconds per second, or 60 watt-microseconds per microsecond, or 60 watt-centuries per century. However, watts do have an embedded reference time unit as part of their joule-based definition: 1 watt equals 1 joule in 1 second (W = J/s). Take note, however, that a joule is the amount of energy required to move an object against a static force of one newton, by the distance that light would travel in 1/299,792,458 second. Consequently, a watt is the amount of power required to push an object against a static force of one newton, at a constant velocity of 1/299,792,458 the vacuum speed of light.

As a measure of the most common unitized form, 1000 units (kilo units), energy is:

- Energy: 1kWH = 1000WH = 1000Wx3600sec = 3600kW−sec = 3600kJ1kWH = 1000WH = 1000Wx3600sec = 3600kW−sec = 3600kJ

21 Units and formulas for: Power

CLARIFICATION: *Joules per second (J/s) is a clearly signified unit of power. Joules per second makes it obvious that power is the rate at which energy is being generated or used. It's like how kilometres per hour (kph) makes it obvious that speed is the rate at which distance is being travelled. Watt as another unit of power; and, as a signifier of [a unit of] power, it does not make it obvious what power means. In other words, the usage of the term 'watt' as a signifier for power, does not make it obvious that power is the rate at which energy is transferred. But, the watt is actually just another name/signifier for Joules per second. J/s and W are the same thing.*

The units of power are units of energy (in a particular system or context) divided by time. The SI unit of power is the watt. The unit of power measurement, the watt, represents energy per unit time. As a rate of change of energy, power is:

- power = change of energy / change of time
- power (P) = ΔE / Δt
- watt (W) = ΔE / Δt

When the rate of energy transfer is constant, power is:

- power = energy transferred / time
- power (P) = E / t
- watt (W) = E / t

The unit of power is joules per second or J/s when work is measured in joules and time in seconds. A watt is the consumption of one joule of energy per second. One watt is equal to one joule of work done per second. Or, said another way, one watt is equivalent to an energy transfer rate of 1 J/second. When energy is measured in joules, then:

- Wattage as J/s = rate of power in Joules per second
- 1 Watt = 1 Joule in 1 second
- 1 W = 1J / 1s
- W = J/s

When energy is measured in newton-meters, then:

- Wattage as Nm/s = rate of power in Newton-meters per second
- 1 Watt = 1 newton-meter in 1 second
- 1W = 1Nm / 1s
- W = Nm/s

Take note that a joule is another term for a force of 1 newton over a distance of 1 meter:

- Joule = Force (1 Newton) x distance (1 Meter)
- 1 Watt = 1J/s = 1Nm/s

As a measure of the most common unitized form, 1000 units (kilo units), power is:

- Power: 1000 Watts = 10^3W = 1kW = 1000J/s = 1kJ/s

21.1 Comparing energy and power in units and formula

Energy is an amount (i.e., quantity), while power is a rate at which energy is used.

- Energy = Watt (Power) x time
  - E.g., kWh = kW x t
- Power = Work (Energy) / Time
  - E.g., kW = kWh / t
- Time = Energy / Power
  - E.g., t = kWh / kW

In concern to units of energy and power:

- Energy is measured in watt-hours (W·h) or joules (J).
- Power is measured in watts (W) or joules per second (J/s).

Watts may be used for [at least] the following power measurements:

- Watts are used to measure the output of a power generating system.
- Watts are used to measure the power production capacity of a power generating system.
- Watts are used to measure the amount of power required by a power consuming system (load).

Watt-time (e.g., watt-hours) may be used for [at least] the following energy measurements:

- Watt-time measures the total amount of energy used over time -- watt-hours is a combination of how fast the energy is used (watts) and the length of time it is used (hours).

Power generating system can be said to produce watt-hours [of energy] per given timeframe (e.g., megawatt-hours per year, and not megawatts per year). Therein, power generating system may be said to produce a specific amount of energy (e.g., Wh) per a given timeframe.

The relationship between energy and power is a lot like the relationship between distance and speed:

- Energy is like distance - The amount of energy that is used over a specific period of time is like the distance that is travelled over a specific period of time. For example, the vehicle travelled 3 meters, or the electrical device used 3 joules (or 3 watt-hours).
- Power is like speed - Instantaneous power is like the speed at a specific instant in time (e.g. right now). The average power over a specific period of time is like the average speed over a specific period of time. For example, the vehicle travelled at a speed of 3 meters per second (m/s), or the electrical device used 3 joules per second (or 3 watts).

22 Power fundamentals

INSIGHT: *From the actualization of potential comes power.*

In the context of power, while 'energy' measures the total amount of energy transferred (i.e., work that is or can be done), it doesn't say how fast the energy is transferred (i.e., how fast the work is or can be done). Herein, power is the rate of transferring, producing, or consuming, energy (i.e., the rate at which energy is transferred, produced, or consumed). Take note that 'power' is not an amount of energy itself; it is a rate of change occurring to the presence of energy. Power is the rate at which a quantity of energy is transferred or otherwise changed in time. It could be the rate at which a quantity of energy is transmitted, as in the case of a power generator, or the rate at which it is received, as in the case of a load. It could also be the rate at which a quantity of energy is transferred between a transmitter and a receiver, such as across a power line between a transmitter and receiver. Thus, power can be described in the following ways, which all amount to the same definition:

1. Power = the [time] rate of energy transfer (or conversion) [within or between energy carriers].
2. Power = the [time] rate of change in energy [in a system].
3. Power = the amount of energy required or expended in a time interval (i.e., for a given amount of time).
4. Power = energy flow per unit time. In other words, power is the time rate of an energy flow. Power can be modelled as an energy flow, equivalent to the rate of change of the energy in a system(s) per period of time.
5. Power = energy "produced", "transformed", or "consumed" per amount of time (i.e., per time interval).

Take note that it is sometimes said that power is "a rate of *energy* generation (production) or consumption (utilization)". Technically, this is not accurate because energy cannot be generated or consumed, it can only be transferred. However, because power is a rate [of transfer], power can be said to be "generated" and "consumed". Power is generated, in the sense that energy transmission is occurring, and it is consumed, in the sense that an end device (load) is using it to function; and that transfer can be started and stopped, slowed or sped up. Hence, although it is not technically accurate to say that "power is a rate of *energy* generation and/or consumption", it is understandable.

DEFINITION: *In mathematics, a 'rate' is the ratio between two related quantities (e.g., A/B, where 'A' is the numerator and 'B' is the denominator). 'Rate' refers to a rate of change. The most common type of rate is "per unit of time" (time denominator), such as speed, heart rate and flux. Ratios (or rates) that have a non-time denominator include: exchange rates; literacy rates; and an electric field (in volts/meter). Often, 'rate' is a synonym of rhythm or frequency, a count per second (i.e., Hertz; e.g., radio frequencies or heart rate or sample rate). 'Power' is a ratio with a time denominator; it is energy per unit time (energy/time = power). Using the signifier "watt" for power is confusing since the "per hour" is not signified by the term itself; instead, it is inside the term "watt". Hence, to make the rate into an amount, it needs to be multiply by a time unit to cancel it out. It would be more intuitive if we worked in joules (energy) and joules per hour or joules/sec (power).*

When the quantity or location of energy changes, then power is present. Therein, as mentioned in the previous paragraph, it may be said that power is the rate at which energy is produced ("generated"), transferred ("transformed"), or used ("consumed") in a given amount of time. For instance, power is the rate at which an system (e.g., electrical) can produce, use, or transfer [electrical] energy. Hence, there are three basic power processes:

1. Power is produced or converted - how much energy is a system producing ("generating" or "transforming") or converting per time interval (e.g., second)? Or, how much power is a system producing from it source of power (energy)?
2. Power is transmitted - how much energy is a system transferring (or delivering) per time interval (e.g., second)? Or, how much power is a system transferring from its source to its usage point?
3. Power is used or dissipated - how much energy is a system using ("consuming") or dissipating per time interval (e.g., second)? How much power is a system using?

In every context, power includes a parameter for 'effort' (or energy) and for 'rate' (time). It is essential to recognize that power is a rate -- a time rate. Thus, regardless of the transfer process, the faster the transfer occurs, the more power is produced. A small amount of energy used extremely quickly can have a lot of power. Similarly, a large amount of energy used very slowly could have very low power. High energy does not necessarily mean high power. The power of any given energy transfer process depends on the time-rate (i.e., "how quickly") a given amount of energy can be transferred. The more energy transferred per time (e.g., /seconds), the greater the power of the transfer.

NOTE: *Power is a widely used measurement. When people speak of the loudness, volume, or level of a signal, they probably mean its power.*

Power is always delivered through pressure (force) and flow [rate] (speed or velocity). Hence, in both mechanical and electrical systems, power delivered may sometimes be calculated by multiplying pressure (force) times flow (speed or velocity). Herein, the rate is included in the flow (speed or velocity) measurement. In mechanical power systems, many terms describe the pressure or force (Newton, Newton per square meter, etc.) and many terms describe the speed or flow (meters per second, litres per second, etc.). In electric power, two terms [at least] describe the pressure or force (voltage and EMF) and two terms describe the speed or flow (current and amperes).

> **INSIGHT:** *All life in the solar system exists because of the power output of the sun.*

Power is absorbed (by a load) and/or transferred (to a load). Power may be dissipated. Power dissipation is the amount of energy per given time period emitted to the outside world by something. In physics, dissipation embodies the concept of a dynamical system where important mechanical modes, such as waves or oscillations, lose energy over time, typically due to the action of friction or turbulence. The lost energy is converted into heat, raising the temperature of the system. Such systems are called dissipative systems. A 60 watt light bulb "dissipates" ~60 watts of power.

There is an upper limit to how much power a power generating system (i.e., energy transformer or transceiver) can output. For instance, a 10 kW wind turbine (provided it has the optimum level of wind), can generate a maximum of 10 kW of power. Hence, 10 kW is the rate at which the wind turbine can generate power, and not the amount of energy that it can generate in a certain period of time. Frequently, the upper limit power output of a power generating system is simply referred to as its "output".

It is frequently said that an electrical device is a device that uses electrical energy. However, such devices are actually transferring the energy to other carriers (i.e., "converting it to other forms") such as heat, motion, electromagnetic radiation, etc., and in the process they are performing a useful function. The rate at which these devices "use" energy is their power [rating]. Depending upon the device and the context in which the power rating is being described, the terms 'load' and 'demand' are synonyms for power [rating]. Take note, however, that while the term 'power' can refer to the power that something is using or generating, the terms 'load' and 'demand' only ever refer to the power that something is using.

> **NOTE: Wattage** *is the maximum power drawn by a device.*

In concern to measurement, electrical devices "use" electrical power measured in units of watts or joules per second (i.e., are powered by watts). As a measure, the watts aren't affected by how long the device is running: a second, an hour, a day - no difference - as long as it's switched on it will be using a certain number of watts of power. If it's not switched on it won't be using any power (i.e., 0 W).

> **NOTE:** *A heat signature always indicates the presence of power (i.e., the presence of the time rate transfer of energy).*

Take note that for some devices it is more complicated to determine power usage/demand. For instance, the watts of power used by a laptop or other computing device may vary from moment to moment depending upon what the system is doing (e.g., how many programs are running). It may be using 50 W of power one moment, 30 W of power the next, and then 43 W of power the next. Hence, the need for a distinction between instantaneous power and average power.

> **HISTORICAL NOTE:** *Why is 'watt' a signifier for a unit of power? For equations, it is simpler for power to have its own unit (instead of being expressed using units of energy and time together). However, some idiot decided to name it after James Watt, the Scottish inventor who facilitated the development of the steam engine, with no relation to earlier energy-associated signifiers.*

Power can be measured (or calculated) in several ways. It can be measured at any instant in time, it can be averaged over a time interval, and its maximum value over a time interval can be determined:

1. **Instantaneous power** (P_i) - Instantaneous power is the power measured at a given instant in time. The instantaneous power (or instantaneous demand, or instantaneous load) is the power that something is using (or generating) at any one moment in time. For example, a 60W light bulb uses 60W every second, and a 60W power generating source can generate a maximum of 60W every second.
2. **Average power** (P_{avg}) - average power is the power measured over a long period (i.e., when t in the equation for power is very large). This is simply the mean, average of the instantaneous power over a longer period of time. Average power may also be referred to as "average load" ("mean load"), or "average demand" ("mean demand"). The average power represents the power that something uses or generates, on average:
   A. Over a specific period of time (e.g. yesterday); or
   B. Over multiple periods of time (e.g. across all the weekends on record); or
   C. Throughout a certain type of operation (e.g. typical laptop usage, or typical building usage - Monday to Friday 09:00 to 17:00, or typical efficiency for something that's generating power).
3. **Peak power** (P_{pk}) - Peak power is the maximum

value the instantaneous power can have in a particular system over a long period.

22.1 Power modes

Every energy transfer mode (or power generation/ consumption system) has its equivalent calculation for power.

1. **Mechanical power** is the rate at which mechanical work is done. Mechanical power is the rate of change of mechanical energy. Mechanical power is the rate at which mechanical energy is converted. In a mechanistic sense, 'power' refers to how far an object can be moved in a given period of time, and hence, how much energy is transferred in that time period.
2. **Electrical power** is the rate at which electrical work is done. Electrical power is the rate of change of electrical energy. Electric power is the rate at which electrical energy is converted. Electrical power is the rate at which electrical energy is "produced" or "used" (or "consumed").
    - Since current is the rate of transport of charge, electric power is given by the above expression, but using current I instead of charge Q:
    - P = IV
3. **Electromagnetic power** is the rate of work done by the electromagnetic forces. Electromagnetic power is the rate of change of electromagnetic energy (Read: electric and magnetic fields).

22.1.1 Mechanical power mode (Work transfer)

In concern to work transfer mode, power is the amount of work done (or, can be done) per unit of time. Power is the time rate at which work is done -- power is the rate of doing/performing work. In other words, power is the amount of work that can be done in a certain amount of time, "the rate of working". Power is the rate of energy transfer by [doing] work per unit of time. Power is work over the amount of time it took to do that work. Regardless of the work being done, the faster the work is done, the more power is produced.

1. Working faster = more power
2. Working slower = less power

Whereas energy is the total amount of work that is or can be done, power is how fast the work is or can be done. Power is also often thought as the amount of work performed (or energy transmitted) in time.

Power is work (energy) per unit of time. Thus, as a rate of change of work done, power is:

- power = change of work / change of time
- power (P) = $\Delta W/\Delta t = \Delta E/\Delta t$
- where, work (W) and time (t)
- power = work done (J) / time (s) = energy used (J) / time (s).

When the rate of work is constant:

- power = work / time
- power (P) = W/t = E/t

22.1.2 Principal types of mechanical working power

There are two principal types of mechancal power, solid and fluid:

1. Mechanical power system (solid mechanics) - linear or rotational motion. Mechanical power systems are used for the generation, control, and transmission of power by the use of solid mechanical objects.
    A. Linear mechanical systems produce linear motion.
    B. Rotational mechanical systems produce angular motion.
2. Fluid power system (fluid mechanics) - linear or rotational motion. Hydraulics are used for the generation, control and transmission of power by the use of pressurized liquids. Pneumatics are used for the generation, control, and transmission of power by the use of pressurized gases.
    A. A hydraulic cylinder or pneumatic cylinder, provides force in a linear fashion.
    B. A hydraulic motor or pneumatic motor, provides continuous rotational motion or torque.
    C. A rotary actuator provides rotational motion of less than 360 degrees.

22.1.3 Linear working power

The full decomposition of the power formula for linear [solid] mechanical work transfer mode has the following sub-parts:

1. In [linear] work (W) transfer mode, energy (E) has the units of force times distance or displacement (F • d). In other words, linear work is force (F) times distance (d). Therein, power is force times distance over time:
    - P = E/t = W/t = (F•d)/t
    - Linear work = Force x Distance
    - W = F•d
    - power = linear work / time
    - P = W/t = F•d/t
    - where, work is in watts, force is in newton, and distance is in meters.

2. Force (F) is mass (m) times acceleration (a) is (m•a):
   - P = E/t = W/t = (F•d)/t = (m•a•d)/t

3. Acceleration (a) is an exponential increase in distance over time (d/t^2):
   - P = E/t = W/t = (F•d)/t = (m•a•d)/t = (m•(d/t^2)•d)/t = (m•d^2)/t^3

4. Alternatively to #2, distance (d) over time (t) is velocity (v): (note that speed is another word for velocity)
   - P = F•d/t = Fv
   - where, velocity is a length measurement per time (e.g., meters per second). The rate is included in the velocity measurement.
   - Note: (1) there must be both force and velocity, and (2) the force must be applied in the direction of the velocity. A static force without velocity does not require power to maintain itself, and velocity (including rotational velocity) without force also does not require power.
   - Note: in electrical systems, power = VI, which is the equivalent in mechanical systems to P = Fv. Therein, voltage (V; a.k.a., electromotive force, EMF) is the force (F), and current (I; a.k.a., amperage) is the velocity (v) .

22.1.4 Rotational working power

The full decomposition of the power formula for rotational [solid] mechanical work transfer mode has the following sub-parts. Note that torque causes objects to spin or rotate.

1. In [rotational] work (W) transfer mode, energy (E) has the units of force (F) times the swept angle (ϴ) times the radius (r).
   - P = E/t = W/t = Fϴr/t

2. Therein, force (F) times radius (r) is torque (T). Hence, energy (E) has the units of torque times the swept angle in radians (T • ϴ).
   - P = E/t = W/t = Fϴr = T•ϴ/t
   - Rotational work = torque x the swept angle
   - W = T•ϴ
   - Power = rotational work / time
   - P = W/t = T•ϴ/t
   - where, work is in watts, torque is in newton-meters, and swept angle is in radians.

3. Swept angle (ϴ) over time (t) is equivalent to angular velocity (Ѡ).
   - P = E/t = W/t = Fϴr =T•ϴ/t = T•Ѡ

22.1.5 Electrical power mode (Electrical transfer)

> **CLARIFICATION:** *The symbol for amperes is generally the letter 'I' (capital 'i'). Before being named amps, the letter 'I' traditionally stood for the "Intensity of current flow", hence, the first letter of the word intensity (I).*

Electric power is the rate, per unit time, at which electrical energy is transferred by an electric circuit (as electricity). In other words, electrical power is the rate of change of electrical charge (electrical energy). The electrical power drawn by an electrical device is expressed in Watts or Volt-Amps (VA). Electrical systems "draw" watts of power. Electric power systems are used for the generation, control, and transmission of power by the use of electricity.

> **NOTE:** *Many electrical devices that dissipate power, do so by converting the electrical power into thermal energy, or heat. This is true for wires and resistors.*

As a rate of doing electrical work in one direction (e.g., DC voltage), power is:

- power = volts x coulombs / time = volts x amps
- power (P) = VQ/t = VI
- where, electric current (I) consisting of a charge of coulombs (Q) every t seconds passing through an electric potential (voltage) difference of V.
- Note, the time interval required for the calculation of power (P=E/t), is included in the electric current parameter (I). Electric current (I) is measured in ampere, and one coulomb per second is equivalent to one ampere (1I).
- Herein, electrical work is done by a voltage (v) moving an amount of electrical charge (q).
  - Electrical work done = voltage difference x charge
  - W = ΔV • q
  - where, the rate of movement of electrical charge is current (I), measured in amperes (I = q/t).

Technically, power in electrical conduction is not transported through electrons pushing length ways just like water in a pipe. It is transported by electro magnetic fields which flow partly inside, partly outside the wire. Electrical energy flows wherever electric and magnetic fields exist together and fluctuate in the same place. The simplest example of this is in electrical circuits, as the preceding section showed. In the general case, however, the simple equation P = IV must be replaced by a more complex calculation, the integral of the cross-product of the electrical and magnetic field vectors over a specified area, thus:

- power (P) = $\int_S (E \times H) \cdot dA$
- where, the result is a scalar since it is the surface integral of the Poynting vector.

A system may take in power (power input) or put out power (power output):

- Input power (power in[put], P_{in}) = voltage x current
- Output power (power out[put], P_{out}) = voltage x current

22.1.6 DC voltage electrical power

In DC circuits, all voltages and all currents are constant, which makes calculation of power simple. A watt is defined as a current of one ampere, pushed by a voltage of one volt.

- Wattage = Volts x Amps
- 1 Watt = 1 Volt x 1 Amp
- 1W = 1V • 1I
- W = VI

In a resistor, the current, voltage and resistance are related by Ohms law as:

- Voltage = Amps x Resistance
- 1 Volt = 1 Amp x 1Ohm
- 1V = 1I • 1R
- V = IR
- where, V is voltage in volts, I is current in amps (traditionally stood for Intensity of current flow), and R is resistance in ohms.

Therein, the power dissipated by a resistor is:

- P = VI
- where, P is power in watts.

Take note that resistors are often rated in both ohms and watts. For a circuit with a single DC power supply, and a single resistor, the power dissipated by the resistor can be written as any of the following forms:

- $P = VI = I^2R = V^2/R$
- where, P is power in watts.

Power cannot be radiated without accelerated charges (i.e. time varying currents). Direct current is time invariant and cannot radiate power.

22.1.7 AC voltage electrical power

AC voltage has phases, and the number of phases may change how power is calculated:

- Single phase, P = IV
- Dual phase, P = IV
- Three phase, P = IV • 1.732
- where, 1.732 is the square root of 3.

22.1.8 Electromagnetic power mode (Electromagnetic transfer)

The watt specifies the rate at which electromagnetic energy is radiated, absorbed, or dissipated.

23 Fundamentals of: Force and motion

NOTE: *To precisely describe motion, the position of an object must be located within a given reference frame. When we say space is three dimensional, we mean we need three numbers to completely locate the position of an object or point. A system for assigning these three numbers, or coordinates, to the location of a point in a reference frame is called a coordinate system. Most frequently, a Cartesian (rectangular) system is used to describes the position in terms of x, y, z coordinates.*

Force is the ability to transfer energy (e.g., a push or pull, a pressure). It is frequently said that a force is a push or pull that one body exerts on another. Thus, a force is always an [inter]action, an influence. It represents the interaction of one body with another, which may be recognized by actual contact or by action at a distance. It is the influence of that which is a 'force' (or 'torque') that produces a change in a physical quantity (i.e., on an object or in a system). When forces are balanced they are said to be in a state of equilibrium. Force is a vector quantity - it has magnitude and direction. Hence, a force is has the following parameters (i.e., is characterized by its):

1. Magnitude
2. Point of action
3. Direction

NOTE: *Vector quantities are often represented by scaled vector diagrams. Vector diagram represents a vector by use of an arrow drawn to scale in a specific direction. Observe that there are several characteristics of this diagram: (1) a scale is clearly listed. (2) an arrow is drawn in a specified direction, therefore, the vector has a head and a tail. (3) the magnitude and direction of the vector are clearly labeled (the magnitude is 100 N and the direction is 35 degrees).*

It is commonly said that force is the "ability to do work". It must be noted that a force is required to do work, but every force does not necessarily do work. To apply a force, an amount of energy is required. This energy is then transferred to the object upon which the force has acted. This force does work on the second object. In this sense, force is a method to transfer energy, thus affecting the motion of a secondary object or system.

Newton's second law of motion states that a force, acting on an object, will change its velocity by changing either:

1. Its speed,
2. Its direction,
3. Or, both.

There are three principle types of motion due to force (i.e., all motion can be classified into three basic types):

1. **Translational/linear motion:** Object moves in a straight line. Translational motion is the motion by which a body shifts from one point in space to another. An object has a rectilinear motion when it moves along a straight line. Translational/linear motion is affected by force. Force causes linear acceleration. Note that "to translate" is "to have linear motion".
2. **Rotational motion:** Object spins. Rotational motion is the motion by which a body rotates in space. Rotational motion is affected by torque. Torque causes angular acceleration.
3. **Vibrational motion:** Object oscillates. Vibrational motion is the motion by which a body moves backwards and forwards (oscillates in two or more degrees of freedom) in space. Vibrational motion is affected by waves (compression and rarefaction). Waves cause vibrational acceleration.

An object can have any combination of these types of motion. For instance, the earth translates around the sun in an elliptical path, rotates about its axis, and vibrates during an earthquake. And, the three types of motion can be separated and analyzed.

In physics, a force is an influencing interaction that causes an object of mass (or charge) to change its velocity. Force can be categorically understood in relation to its physical application.

1. For a mechanical system, when force is applied, mass is displaced, and work is done (energy is transferred). Or, when torque is applied, mass is rotated, and work is done (energy is transferred). Therein, power is present as linear or rotational movement occurs.
   - Work (W) = force (F) x displacement (d)
   - Work (W) = $F \cdot r\Theta = \tau\Theta$
2. For an electrical system, when force is applied, charges flow, and work is done (energy is transferred). Therein, power is present as charges flow.
3. For an electromagnetic system, when force is applied, electromagnetic waves (perturbations) propagate, and work is done (energy is transferred).

A force can cause any of the three types of motion. Therein, it could be said that there is one principal type of motion instantiation, and one principal sub-type of motion instantiation. The principal type [of motion] is force, and the principal sub-type [of motion], which is caused by the principal type (force) is torque.

1. [Principal type] **Force** as that which causes linear

(translational) motion. Translational motion is affected by force.

2. [Sub-type] **Torque** (twisting force, moment, moment of force) as that which causes rotational motion. Rotational motion is affected by torque. Torque is a measure of how much a force acting on an object causes that object to rotate. Torque is the counterpart of the force in angular motion.

A force can be acted as a force alone or as a torque. A force can be present without a torque, but a torque cannot be present without a force. A force is necessary in order to create a torque. Torque is created by a force. The specifics of the torque depend on the location of the force and the center of rotation (i.e., point about which an object rotates, the pivot point). One important distinction between force and torque is direction. Positive and negative signs are used to represent forces in the two possible directions along a line. The direction of a torque, however, is clockwise or counterclockwise, not a linear direction.

Take note that it is possible [for an object] to have a zero total torque with a non-zero total force. For instance, an airplane with four propeller engines -- two on either side of the fuselage, each side's propellers spinning in opposite directions to cancel out the total torque. Conversely, it is possible to have a zero total force and non-zero total torque. A merry-go-round's engine(s) need to supply a non-zero torque to bring the go-round object up to speed, but there is a zero total force on it. If there was not zero total force on it, its center of mass would accelerate.

NOTE: *The farther away from the center of rotation that the force is applied, the easier it is to rotate, the greater the torque.*

A simplistic way of classifying force is as follows:

1. **Applied force** is a push or a pull that is exerted on an object by another.
2. **Force of gravity** is the natural force that draws any object that is thrown to the sky towards the center of the earth.
3. **Normal force** is the magnitude of push that is brought about by an object's own weight.

A force is the fundamental result of an interaction between two objects, whereas power is an expression of the rate of energy transmitted over time (e.g., work), of which force is an element. Force and power can both be described and measured, but a force is an actual physical phenomenon, and power in itself is not.

NOTE: *Heat transfer by friction involves force. Heat transfer by conduction does not involve force. The definition of work could thus be restated as the amount of energy transferred by forces. No work is done without motion.*

23.1 Mechanical force

Mechanical force includes several possible sub-types of force.

23.1.1 Linear motion (linear/translational force)

Sum of all forces (f) = mass (m) x acceleration (a). Mass times acceleration (ma) is not a force. The sum of all forces on an object equals the product of its mass times its acceleration.

If every part of a system moves in a straight line at a constant speed, then it is in translational equilibrium (Note: this includes being at rest). For a body to be in translational equilibrium, the resultant forces in any two perpendicular directions must be zero.

- Force = Mass x Acceleration
- Newtons = kg x m/s^2
- where, the acceleration of gravity on earth is 10 m/s^2.

23.1.2 Torque (rotational force)

A force that produces a twisting or turning effect, or rotation, is called torque. Torque is also called a "rotational force" or a "twisting force". It is a "force" that makes anything rotate, twist, or turn. Any time anything rotates, there is a torque involved. Torque is the rotational equivalent of linear force. Torque can be used to create a force at a distance, but it does not cause an object (directly) to move along a distance. Torque is defined as a force around a given point (axis), applied at a radius from that point. A force applied at a non-zero distance from an object's centre will tend to rotate the object. This is easily seen in real life. If a wrench is placed on a bolt and a force is applied to the end of the wrench, the bolt will turn. If the same pulling force was applied directly to the bolt, it would not turn because the force's direction passes through the object's centre. The amount of torque is determined by multiplying the magnitude of the force by the force's distance from centre.

The ability of a force to rotate an object about an axis depends on two variables:

1. The magnitude of the force (F).
2. The distance (r) between the axis of rotation and the point where the force is applied.

The "turning ability" of a force is the product of F and r. The technical name for this "turning ability" is torque. Hence, the torque τ exerted by a force F that is applied at a point r relative to the origin is the cross product of r and F. Thus, the formula for torque is:

- torque (τ) = force (F) x perpendicular distance (r)
- $\tau = F \cdot r$

The magnitude of a torque depends on three quantities:

1. The force applied.
2. The length of the element (e.g., lever arm) connecting the axis to the point of force application.
3. The angle between the force vector and the lever arm.

Note that the units for both torque and work are the product of force and distance, yet torque and work are two different things. Torque is a force that tends to cause a rotation, which means that it does not actually cause an object to move along a distance. Work is a measure of energy transfer between systems, which may or may not have been done by a force from torque.

Mathematically, for rotation about a fixed axis through the center of mass, the formula is:

- $W = \int^1_2 \tau d\theta$
- where, W is work, τ is torque, and ∫ (θ1 and θ2) represent (respectively) the initial and final angular positions of the body.

Whereas torque is measured, power is calculated. The power of a torque (rotational force) is a product of torque and rotational speed (i.e. cadence). The power (work per unit time) of a torque is given by:

- power (Pm) = torque (τ) · angular velocity (Ѡ)
- where, Ѡ is angular velocity or angular speed.
- Note, power herein is notated as a mechanical parameter, hence, mechanical power (Pm).

The terms “moment” and “torque” are often used interchangeably. By definition, however, moment is a quantity that represents the magnitude of force applied to a rotational system at a distance from the axis of rotation.

23.1.3 Pressure

Pressure is the force applied perpendicular to the surface of an object per unit area over which that force is distributed. Pressure is the ratio of force to area over which it is applied. Pressure is a scalar quantity as it has magnitude but no direction, while force is a vector quantity, because it has both magnitude and direction.

- pressure (P) = force (F) / area (A)
- F/A = ΔF/ΔA = work/volume = energy / volume

Pressure in a fluid can be seen to be a measure of energy per unit volume by means of the definition of work. This energy is related to other forms of fluid energy by the Bernoulli equation. Pressure in a fluid may be considered to be a measure of energy per unit volume or energy density. For a force exerted on a fluid, this can be seen from the definition of pressure:

23.2 Electrical force

When two bodies of matter have charges and are near one another, an electric force (F) is exerted between them. The existence of such force, where current does not flow, is referred to as static.

The force of attraction or repulsion exerted between two charged bodies is directly proportional to the product of their charges (Q) and inversely proportional to the square of the distance (d) between them.

This relationship between attracting or repelling charged bodies was first discovered by a French scientist named Charles Coulomb and accordingly is known as Coulomb’s Law:

- $F = \hat{a} (Q_1 Q_2 / 4\pi\varepsilon_0 d^2)$
- Electric force (F) = $k (q_1 q_2 / r^2)$
- where, F is a vector quantity, which represents the electrical force acting on charge Q_2 due to charge Q_1 measured in newtons (N).
  - Wherein, F is the electric force.
- where, â is a dimensionless unit vector with a unity magnitude pointing from charge Q_1 to charge Q_2. k is coulomb's constant.
  - Wherein, q1 and q2 are charges (scalar values).
- where, ε_0 is a universal constant called the electrical permittivity of free space [$\varepsilon_0 = 8.854 \times 10^{-12}$ farad per meter (F/m)].
  - Wherein, r is the distance of separation between the two charges.

24 Fundamentals of: Electricity

Electricity is the set of physical phenomena associated with the presence and flow of electric charge. Electricity gives a wide variety of well-known effects, such as lightning, static electricity, electromagnetic induction and electric current. In addition, electricity permits the creation and reception of electromagnetic radiation such as radio waves.

In electricity, charges produce electromagnetic fields which act on other charges. Electricity occurs due to several types of physics:

1. **Electric charge:** a property of some subatomic particles, which determines their electromagnetic interactions. Electrically charged matter is influenced by, and produces, electromagnetic fields, electric charges can be positive or negative. Electric charge is measured in coulombs.
2. **Electric force:** the force of attraction or repulsion between objects due to charge.
3. **Electric field:** charges are surrounded by an electric field. The electric field produces a force on other charges. Changes in the electric field travel at the speed of light. Electric fields are measured in kilovolts per metre (kV/m).
4. **Magnetic field:** these fields are toroidal/spiral in form, near instantaneous, and feedback into/onto themselves. They are not as well studied as electrical fields. A changing/fluctuating (in flux) magnetic field will induce an electric field [at a distance] in a material capable of conducting an electric field (i.e., in a conductive material). A magnetic field can be generated in one of two ways (either / or): directly through a permanent magnet; or indirectly by passing an electric current through conductive coils/windings to produce an 'electromagnet'. and Magnetic fields are measured in milligauss (mG).
5. **Electromagnetic field:** The synchronized and perpendicular propagation of electric and magnetic fields at a specific frequency. Electromagnetic field are typically measured in electron-volts (eV). The frequency is typically measured in hertz (Hz).
6. **Electric potential:** the capacity of an electric field to do work on an electric charge, typically measured in volts.
7. **Electric current:** a movement or flow of electrically charged particles, typically measured in amperes.
8. **Electric conductor**: a material that can carry an electrical current, and through which charges move freely. Conductivity is determined by the atomic makeup of a material. Materials with high electric charge mobility (many free electrons) are called conductors, while materials with low electron mobility (few or no free electrons) are called insulators. For electrons to flow continuously (indefinitely) through a conductor, there must be a complete, unbroken path for them to move both into and out of that conductor. Note that when a current carrying conductor is placed in a magnetic field it experiences a force (specifically, an electromagnetic force).
9. **Electromagnets:** moving charges produce a magnetic field. Electric currents generate magnetic fields, and changing magnetic fields generate electric currents.

In electrical engineering, electricity is used for:

1. **Electric power** where electric current is used to energise equipment.
2. **Electronics** which deals with electrical circuits that involve active electrical components such as vacuum tubes, transistors, diodes and integrated circuits, and associated passive interconnection technologies.
3. The **electromagnetic spectrum** is the range of all possible frequencies of Electromagnetic radiation and is sorted by frequency. The electromagnetic spectrum is a categorized spectral representation of electromagnetic waves by their wavelength (frequency) location within the whole set of known electromagnetic waves, which is most commonly measured in micrometers. Names are often assigned to regions of the electromagnetic spectrum, but there is no clear cut dividing lines from one region to the next. Electromagnetic waves can be characterized by either the frequency or wavelength of their oscillations, which determines their position in the electromagnetic spectrum, which includes, in order of increasing frequency and decreasing wavelength. The electromagnetic spectrum is a unified spectrum of photonic energy patterning that humans, for purposes of functional service specialization, have split into "frequency bands". Individual photon energies in the frequency band known as "radio frequency" (RF) are so small that its not useful to describe RF waves in terms of photons, but one could do so. Each "band" represents electromagnetic radiation along a continuous spectrum, split into wave segments with different, upper and lower, frequency boundaries (a.k.a., "bands"). So, visible light and RF are the same thing, EMR; its only a matter of how much energy each of them (i.e., each frequency instantiation) is carrying. Because of its frequency, the "RF" band is has lower photon energy than visible light.

24.1 Electricity in nature

Electricity is not a human invention, and may be observed in several forms in nature, a prominent manifestation of which is lightning. Many interactions familiar at the macroscopic level, such as touch, friction and chemical bonding, are due to interactions between electric fields on the atomic scale. The Earth's magnetic field is thought to arise from a natural dynamo of circulating currents in the planet's core. Certain crystals, such as quartz, and even sugar, generate a potential difference across their faces when subjected to external pressure. This phenomenon is known as piezoelectricity, and was discovered in 1880. The prefix "piezo-" is derived from the Greek piezein (πιέζειν), which means to press or to squeeze. The effect is reciprocal, and when a piezoelectric material is subjected to an electric field, a small change in physical dimensions takes place.

Some organisms, such as sharks, are able to detect and respond to changes in surrounding electric fields, an ability known as electroreception, while the ability of an organism to internally generate an electric voltage is termed electrogenic (such an ability often serves as a predatory or defensive weapon). The biological order Gymnotiformes, of which the best known example is the electric eel, is able to stun its prey via high voltages generated from modified muscle cells called electrocytes. It is important to note here that all animals transmit information along their cell membranes with voltage pulses called 'action potentials'. Action potentials are also responsible for coordinating activities in certain plants.

24.2 Principles of electrical theory

The following principles form the foundation of electrical theory, and hence, are the basis of electricity.

1. All charged particles (i.e., charges) have an electric field. There are no magnetically charged particles (if they existed they'd be called "magnetic monopoles").
2. There are two ways of creating an electrical field: introduce an electrically charged particle, or introduce a time-varying magnetic field.
3. There are two ways of creating a magnetic field: move an electrically charged particle (i.e., an electrical current), or introduce a time-varying electric field.
4. Under static conditions (i.e., not changing with respect to time), either electric or magnetic fields can exist without the other. Technically, this is not entirely accurate in concern to magnetic fields, because the electric field is still present, it is just being cancelled out.
5. Under dynamic conditions (i.e., changing with time), neither an electric or magnetic field can exist without the other. If one field is time-changing, the other must be non-zero.
6. Electromagnetic energy refers to the synchronous presence of electric and magnetic fields.
7. Electrostatics is the study of static (unchanging) electric fields, electric charges, and the rules governing their interactions.
8. Magnetism is the study of static magnetic fields, magnets, and the rules for their interactions.
9. These two areas of study are tied together with electrodynamics, which is the study of changing electric and magnetic fields, and electromagnetic (EM) waves (of propagation radiating as "radiation").
10. Electrical science is the study of electrical effects. Electrical effects are caused by electric charges and by the electric and magnetic fields associated with charges.
11. The theory of electric circuits is a subset of the theory of electrodynamics, which is a subset of quantum theory.

INSIGHT: *There is no physical object called 'wave'. Wave is not what something is, but what something does. For instance, wave is what a flag does. There is no waving without the flag. Similarly, there is no physical object called vortex. Vortexing is what something does.*

24.3 Electric charge

The 'electric charge' is a property of some subatomic particles, which determines their electromagnetic interactions. Charge is the quantity of electricity responsible for electric phenomena. It is one of the fundamental physical quantities in electric circuit analysis. A quantity of charge that does not change with time is typically represented by Q. The instantaneous amount of charge, which is time dependent (changes over time), is commonly represented by Q(t). The concept of an "electric charge" adheres to the following principles:

1. The physical property of matter that causes it to experience a force when placed in an electromagnetic field is called an "electric charge", and it has historically been called a "charge of electricity".
2. Electric current is a "flow of charge", rather than "a flow of electricity."
3. Electrons are "charge carriers", rather than "particles of electricity."

CLARIFICATION: *The speed at which energy transfer occurs, or signals travel, down a conductor is the speed of the electromagnetic energy (light), not the speed of movement of the electrons.*

The decomposition of the physical environment to charge is as follows:

1. Matter - all forms of matter are composed of molecules.
2. Molecule - molecules are composed of atoms
3. Atom - atoms are composed of particles, of which there are three: protons, neutrons, and electrons.
    A. Relative position - the center of the atom is called the nucleus, which is surrounded by orbits.
    B. Protons within nucleus.
    C. Neutrons within nucleus.
    D. Electrons in orbit.
4. Charge - a property of particles determining electromagnetic interaction.
    A. Electrons maintain a negative polarity (-ve), a negative charge.
    B. Protons maintain a positive polarity (+ve), a positive charge.
    C. Neutrons do not have a polarity, are neutral.
5. Atomic charge determination
    A. An excess of electrons creates a negative charge.
    B. The absence of electrons creates a positive charge.
6. Charge field vector
    A. Positive charge outward
    B. Negative charge inward
7. Charge interaction/dynamics
    A. Different charges attract each other (void space).
    B. Same charges repel each other (create space).

In physics, 'charge', also known as electric charge, electrical charge, or electrostatic charge (dielectric), and symbolized 'q', is a characteristic of a unit of matter that expresses the extent to which it has "more (-ion) or fewer (+ion) electrons than protons". The basic unit of electric charge is the "coulomb".

- 1 coloumb = 6.25 x 10^{18} electrons

Electric charge, also called "the Quantity of Electricity," is a fundamental component of everyday matter. Objects are made of molecules and atoms, atoms are made of protons, neutrons, and electrons, and the protons and electrons are made in part out of electric charge. Electric charge is substance-like. If you have a quantity of charge, you cannot destroy it, you can only move it from place to place.

> **NOTE:** *Electric forces are what hold together atoms and molecules, solids and liquids. In collisions between objects, electric forces push things apart.*

Matter can carry charge. However, it is not the charge of matter that transports energy; it's the electromagnetic field that is linked to the charge. Charged particles are expressed as propagating electromagnetic excitations in the field. In a charged particle's rest frame (static charged), the magnetic components is not expressed, and only the time-like ones (the electric field and the energy, respectively) remain. Charge itself gives rise to a 'divergence' in the electric field. Current (moving charge) gives rise to a curl/spiral in the magnetic field. In other words, in its rest frame, a charged particle appears to generate an electric field only and no magnetic field at all. From a different frame of reference (in particular one in relative motion), we'll see the charge moving, thus a current which generates a magnetic field as well. Fundamentally, charge produces a field that acts on other charges.

> **NOTE:** *Materials can be listed in the order of those "most likely to lose electrons" (gaining positive charge) to "those most likely to gain electrons" (gaining negative charge). This is called the 'triboelectric series'.*

Note, it is possible to let charges pass through a vacuum with no resistance, but that is not a reason to call the vacuum a conductor [of charge]. Conducting is associated with influence of the conductor on the motion of the conducted - directing the motion - which a vacuum does not appear to have. The vacuum (space) allows charged matter to pass trough it. Electromagnetic waves propagates (at the speed of light) in vacuum.

24.4 Charge and electric circuits

Moving charges represent an electric current. During operation, although charges are transferred between different parts of an electric circuit, the total amount of charge does not change. Electrons or protons are neither created nor destroyed when an electric circuit is operating.

In a neutral state (zero charge), electrons will neither leave nor enter the neutrally charged body should it come in contact with other neutral bodies. If, however, any number of electrons is removed from the atoms of a body of matter, there will remain more protons than electrons and the whole body of matter will become electrically positive.

Should the positively charged body come in contact with another non-charged body, or having a negative charge, an electric current will flow between them. Electrons will leave the more negative body and enter the positive body. This electron flow will continue until both bodies have equal charges. When two bodies of matter have charges and are near one another, an electric force (F) is exerted between them. The existence of such force, where current does not flow, is referred to as 'static'.

The force of attraction or repulsion exerted between two charged bodies is directly proportional to the product of their charges (Q) and inversely proportional to the square of the distance (d) between them.

24.5 Conductors

Conductors allow for charge transfer through the free movement of electrons (or protons). Conductors guide the flow of electric charge, and hence, the flow of electromagnetic energy.

Three factors determine whether or not the atom is a "good" or "bad" conductor:

1. The number of electrons in the outer orbit.
2. The distance of the outer orbit from the nucleus of the atom.
3. The density of the atoms within the element.

Therein,

1. If the atom has only one orbit, maximum number of electrons on orbit is 2.
2. If the atom has more than one orbit, maximum number of electrons on outer orbit is 8.

The following are good conductors:

1. Gold, silver, copper have 1 electron on their outer orbit.
2. Mercury has 2 electrons in its outer orbit
3. Aluminum has 3 electrons in its outer orbit.
4. Carbon has 4 electrons in its outer orbit.

> **NOTE:** *The net electric charge of a conductor resides entirely on its surface. The mutual repulsion of like charges from Coulomb's Law states that the charges be as far apart as possible, hence on the surface of the conductor. The electric field inside the conductor is zero.*

24.6 Electric current

> **NOTE:** *Electron theory states that the subatomic particle that does the work in electronics is the electron, which happens to be negatively charged. There is a subatomic particle that flows the other way, from positive to negative: the Positron.*

An electric current is a flow of electric charge, which transfers electromagnetic energy through conductive space. The particles that carry the charge in an electric current are called 'charge carriers'. There are a variety of charge carriers:

1. In metallic solids, electric charge flows by means of electrons, from lower to higher electrical potential.
2. In electrolytic solutions, electric charge flows by means of ions.
3. In gases and plasmas, electric charge flows by means of ions and electrons.
4. In a vacuum, electric charge flows by means of ions and injected free electrons.

Electric current is measured in coulombs per second (amperes or amps; A or I).

- Amperage = amount of electrical current
- Amperage = Coulombs / Seconds
- 1 Ampere = 1 Coulomb / 1 Second
- 1 Ampere is equal to 1 Coulombs per second
- 1A = 1C / 1s
- A = C/s

> *1 Coulomb is approximately 6.241×10^{18} times the elementary charge (e or q). The elementary charge is the electric charge carried by a single proton, or equivalently, the magnitude of the electric charge carried by a single electron (-e or -q). This elementary charge is a fundamental physical constant.*

Current is rate of change in the electric field:

- current (I) = $\Delta q / \Delta t$
- wherein, q=charge

> **NOTE:** *Ampère's force law states that there is an attractive or repulsive force between two parallel wires carrying an electric current. This force is used in the formal definition of the ampere, which states that the ampere is the constant current that will produce an attractive force of 2×10^{-7} newtons per metre of length between two straight, parallel conductors of infinite length and negligible circular cross section placed one metre apart in a vacuum.*

Electric current is measured using an instrument called an **ammeter**.

> **NOTE:** *Humanity cannot [with present technology] directly observe the electrically-charged particles that produce current.*

The energy in electric circuits is not carried by individual electrons, it is carried by the circuit as a whole. Current is defined as the rate of flow of charges through a medium. Current is the flow of charges, stationary charges cannot give any current. Charge gives rise only to an electric field, while current produces both electric and magnetic fields. The energy flowing through an electric circuit as an electric current is contained in electrostatic and the magnetic fields produced by the electrons.

> **NOTE:** *In an electric current, the electron particles are the "medium", wherein energy is transferred electromagnetically.*

Electricity (electrical energy) is the flow of electrical charge, and all flows of electrical charge (electric current) transfer energy electromagnetically. In metals, the electrically charged particles are electrons. Electricity cannot flow through air, except in the form of electrically charged particles of air - as in a spark or lightning stroke.

NOTE: *Some elements in a circuit can transfer ("convert") energy from one carrier ("form") to another. For example, a resistor transfers ("converts") electromagnetic energy traveling through a conductor (i.e., "electrical energy") to heat (i.e., "thermal energy"), this is known as the Joule effect. In other words, electric currents cause Joule heating.*

NOTE: *The flow of charge (i.e., an electric current) causes friction, which is called resistance. Resistance quantifies how much current you get across something per volt applied. Namely, if you apply a voltage V across a wire and measure current I, the resistance R is defined by: R = V/I. Resistance therefore has units of V/A, which get another name, ohms.*

24.6.3.1 Current and the AC power grid

In an AC electric power grid, a certain amount of energy is lost because it vectors off into space. This is well understood: electrical energy is electromagnetic waves travelling everywhere, and unless the power lines are twisted or somehow shielded, they will act as 50-60Hz antennas. Waves of 60Hz electrical energy can spread outwards into space rather than follow the wires. The power lines can even receive extra 60Hz energy from space, from magnetic storms in Earth's magnetosphere. Electric energy is gained and lost to empty space while the charges of electricity just sit inside the wires and wiggle.

24.7 Current and electromagnetic fields

Any time current flows through a conductor, a magnetic and electric field are generated around the conductor. If that current is direct current (DC), then the resultant magnetic and magnetic fields will have a constant orientation/polarity (i.e., a constant electromagnetic field - DC magnetic field and DC electric field). If the current is alternating current (AC), then the electric and magnetic fields will vary in direction (polarity) and intensity with the alternation of the current (i.e., a varying electromagnetic field - AC magnetic field and AC electric field).

Any AC circuit propagates its signals using electromagnetic waves. The transmission of the signal between elements is done only by electromagnetic waves. But in a circuit, these waves are guided waves, the traces on a PCB or the wires of our circuit guide the waves along the desired path.

NOTE: *An antenna is a transceiver/transformer of sorts. The antenna is a device that transforms guided electromagnetic waves into propagating electromagnetic waves (and vice-versa). So all it is doing is taking the guided wave that is sent to the antenna and providing it a means of going into open space.*

24.8 Electromagnetic fields

An **electromagnetic field (EMF** or **EM field)** refers to the field created by static (electric field) or moving (magnetic field) charges. Note that a constant electric or magnetic field filling a space is not a wave (i.e., not an electromagnetic wave). In physics, a field is a space together with a set of values for every point in the field, which generates a time-space coordinate system. An electromagnetic field is a set of values for electric and magnetic vector orientation and magnitude (strength), one for each point in space time. The components of the field depend on a reference point for the coordinate system (the observer), even though the field itself has a definite physical existence. Technically, a classical field is a function whose domain is space-time, and a wave is a configuration of the field that satisfies a [differential] wave equation. Note that a quantum field is more complicated. Note that the term 'field' is challenging to define because it is a fundamentally assumed/axiomatic form of existence in physics. Hence, it cannot be defined by saying what it is made of. Electric fields are measured in kilovolts per metre (kV/m) and magnetic fields are measured in milligauss (mG).

An electric field can be created by:

1. The presence of a changing magnetic field.
2. The presence of a charge[ed particle] (e.g., ion).

A magnetic field can be created by:

1. The presence of a changing electric field.
2. The presence of a dielectric charge[d particle] (e.g., permanent magnet).

24.8.1 Alternating current and electromagnetic fields

The electromagnetic fields produced by an alternating current can be categorized as follows:

1. *Near fields* allow for electromagnetic induction. The near-field is a reactive power field. Inductive coupling is the coupling of elements with near fields.
2. *Far fields* allow for electromagnetic radiation.

24.8.1 Alternating current and near field electromagnetic induction

In a coil of wire, AC produces fluctuating fields that can induce currents on another coil without any physical contact. This process is known as electromagnetic induction, and it uses near field electromagnetic reactance radiation (vs. far field propagating radiation). The electric and magnetic fields produced for electromagnetic induction are not in a constant ratio of strengths to each other, and are not phase (i.e., there is a reactance). Electromagnetic induction is a particular form of the more general electromagnetic field (EM Field), which is produced by moving charges. If an AC current is fed through a piece of wire, the electromagnetic field that is produced is constantly growing and shrinking due to the constantly changing current in the wire. This growing and shrinking magnetic field can induce electrical current in another wire that is held close to the first wire. The current in the second wire will also be AC and in fact will look very similar to the current flowing in the first wire. One can generate a magnetic field by letting an alternating current flow through a wire or coil. That is what happens in the primary coil of a 'transformer'. The other way around, a change in a magnetic field will generate a current in a coil - that's what happens in the secondary coil. These properties of magnetic fields and current are called electromagnetic induction.

> **NOTE:** *The near field and far field are regions of the electromagnetic field around an object, such as a transmitting antenna, or the result of radiation scattering off an object. This difference between picking up a magnetic field and magnetic radiation is known as the difference between near and far field. [wikipedia.org]*

In a general electromagnetic induction setup, the secondary coil exists inside one wavelength of the wave that is produced by alternating current on the first coil (i.e., in the near field). This means that the current in the secondary coil does not exist because of electromagnetic radiation (self-propagating EM fields), but because of electromagnetic induction (reactance EM fields). In an electromagnetic induction circuit, the electric and magnetic fields don't [re-]create each other and propagate outward.

24.8.2 Alternating current and far field electromagnetic radiation

Electromagnetic radiation (EMR) is a particular form of the more general electromagnetic field (EM Field), which is produced by moving charges. The electric and magnetic fields in EMR exist in a constant ratio of strengths to each other, and must also be in phase. In electromagnetic radiation, the magnetic field will create an electric field (just assume that), but further away from the conductor that began with making the electromagnetic field. The electric field will create a magnetic field, even further away, and so on. It just goes on and on, due to specific properties of the field. It can vary in frequency, from extremely low frequency all the way up to extremely high frequency.

Electromagnetic radiation (EM radiation or EMR or far field) is the radiant electromagnetic energy released by certain electromagnetic processes. Electromagnetic radiation is a transverse wave where an electric and magnetic field oscillate perpendicular to each other and in the direction of propagation. The energy of the wave is in the electric and magnetic fields. Electromagnetic radiation is associated with those EM waves that are free to propagate themselves ("radiate") without the continuing influence of the moving charges that produced them, because they have achieved sufficient distance from those charges. Thus, EMR is sometimes referred to as the **far field**; versus, the near field, which refers to EM fields near the charges and current that directly produced them, specifically, electromagnetic induction and electrostatic induction phenomena.

> **NOTE:** *In general, electromagnetic radiation from an antenna comes from alternating current flowing in a linear conductor.*

24.8.3 Direct current and electromagnetic fields

A direct current (DC) electromagnetic field refers to a constant or static DC electric or DC magnetic field emission, which has a frequency of 0 Hz. In a coil of wire, DC produces electromagnetism, and does not produce electromagnetic induction (near field) or electromagnetic radiation (far field). A DC magnetic field (constant polarity) cannot be used to induce current in any other conductor. Only a varying magnetic field can do that (to generate that you need varying current). You can use this unidirectional field in a way similar to how you can use permanent magnets. For example - closing and opening electromechanical relays. The only way to produce an electromagnetic field is to somehow change the current with time. So, even if the source of the current is constant (DC), then it is possible to produce an EM field by frequently changing the physical properties of conductor along its length, such as changing the cross-section of the conductor frequently along its length, or modifying the electrical parameters of the conductor frequently. The electric field of a direct current (DC electric field) is measured in are measured in Volts per meter (V/m). The magnetic field of a direct current (DC magnetic field) is measured in milliGauss mG with a DC gaussmeter.

24.9 Electromagnetic radiation

Electromagnetic radiation (in the shape/geometry of a wave) are produced by accelerating electrical charges. The current carrying charged particles in AC circuits are continuously accelerating (at a frequency) and always emit electromagnetic waves. These emissions may be limited to reduce energy losses with the use of shielding, twisting, and coaxial cables. An EM wave is present when there is an oscillation of charge (as in, an oscillator produces a periodic, oscillating electron signal, an AC signal).

In a DC circuit operating with a constant current, the electrical charges, usually electrons, only experience a brief initial acceleration when the circuit is energized, and negligible energy is radiated as electromagnetic waves. In other words, the DC hasn't been DC forever, there was a time when it turned on, and that put out a small electromagnetic click, but just for an instant. A DC wire puts out a steady magnetic field, not a propagating electromagnetic field. Direct current is also capable of producing a varying magnetic field (by turning it on and off at a certain frequency, for example), So, it may emit an electromagnetic wave, if it's varying in some way.

The electric and magnetic fields produced by direct current (DC) lines are referred to as static fields because their sources, voltage and current, do not alternate over time. Thus, DC fields are qualitatively different in nature than the alternating current (AC) electric and magnetic fields (often called EMF) produced by AC transmission lines. While AC EMF can cause the induction of currents or voltages in nearby objects, this does not occur with DC fields.

Stable AC produces a constant "vibration" in the conductor, while DC doesn't vibrate the conductor at all. If the electron flow in 60Hz AC power signal were converted to a sound, then it would sound like a low hum -- specifically, a 60 Hz hum, between a B and a Bb, right below the C two octaves below middle C. DC current sounds like a single click when it starts and another when it ends. This is because what we call sound consists of vibrations.

Electric and magnetic fields surround any electrical circuit, whether it carries AC or DC power, including appliances, electrical wiring and power lines. Both electric and magnetic fields diminish rapidly as the distance from the source increases. Electric and magnetic fields from DC transmission lines are commonly referred to as static fields because they do not alternate in direction. Static electric fields occur as a result of voltage. Static magnetic fields are created by a magnet or by the steady flow of electrical current (DC).

The fields associated with the operation of a DC line are static, which is the same as having a frequency of zero, and do not induce voltages or currents in nearby conducting materials in the environment. Note that in certain weather conditions, both AC and DC transmission lines may produce an electric field associated with electric charges in the air and not just those on the conductors.

An electric field applied to an electric circuit causes a flow of electric charge, which transports/moves electromagnetic energy and generates consequential heat as thermal energy due to resistance (friction). All charges have an electric field. When you accelerate a charge you also get a magnetic field. To get EM waves you need to accelerate the charges - like wiggling them back and forth or turn them in a circle for acceleration. Electrons accelerating in a conductor do emit EM waves out of the conductor - that is how radio transmitters work. A DC circuit does not emit significant EM waves while it transports/moves electrical energy from source to load.

In electronics and telecommunications engineering, a "transmitter" or "radio transmitter" is an electronic device which, with the aid of an antenna, produces EMR (as radio waves). The electronics of the transmitter device generate a radio frequency alternating current, which is applied to a part of the device known as an antenna. When excited by this alternating current, the antenna radiates EMR (as radio waves). The term transmitter is usually limited to equipment that generates radio waves for communication purposes, or radiolocation, such as radar and navigational transmitters. Generators of radio waves for heating or industrial purposes, such as microwave ovens or diathermy equipment, are not usually called transmitters even though they often have similar circuits.

The electromagnetic fields that we measure radiating from AC electric currents in the circuits in the walls of a building have a frequency of about 50 to 100 cycles per second. If we increase the frequency to 500 or 1000 kilocycles (1 kilocycles = 1000 cycles) per second, we are "on the air", for this is the frequency range which is used for radio broadcasts.

All accelerated charges radiate electromagnetic energy (i.e., electromagnetic radiation). So, everything that conducts alternating current acts as an antenna. However, in order to achieve efficient radiation the antenna must be designed appropriately. The antenna itself, when connected to a transmitter, is both the positive, 0, and negative pole at different times. This movement of charge creates a changing electric and magnetic field, which becomes an electromagnetic wave, capable of radiating energy from the antenna or aerial (see Maxwell equations and Hertz definition). As the [alternating] current from the transmitter approaches the end of the wire [antenna], but has no place to go, the charges pile up until they are pushed back in the other direction. By the time the charge is back at the transmitter, it's travelled $\lambda/2$ or experienced a 180° phase shift. The voltage at V1 has also changed by this point, and so the current is constructively adding to the new currents being produced by the transmitter, as an alternating current that form a sine wave. If it were not for some of this energy being "lost" as radiation, the energy in the antenna would grow without bound. The radiation of energy from the antenna is presently understood in the form of a set of equations named after a human being, "Maxwell's equations". Essentially, the equations state that the current in an antenna is associated with a magnetic field, and the voltage is associated with an electric field -- an antenna is an arrangement such that at some distance away from the antenna (the far field) these two fields are mutually perpendicular and in phase, and the output of their integration [in a medium] is a self-propagating EM [field] wave.

INSIGHT: *An equation may be true, but not* *factual.*

The electric field is produced by stationary charges, and the magnetic field by moving charges (currents; or, permanent magnetic substance); these two concepts are often described as the sources of the field. The way in which charges and currents interact with the electromagnetic field is described by Maxwell's equations and the Lorentz force law. The alternating voltage from the transmitter is moving (accelerating) the [electron(ic)] charge backwards and forwards. Standing waves that impact the functioning of the antenna are the result of a miscalculated antenna, and they represent lost energy. The standing wave is the pattern you get (in voltage or current) when the power travelling to the antenna is superimposed on the power reflected back from the antenna due to mismatch of antenna and transmission line. Power is travelling in both directions at once and when you sum the instantaneous voltage at all points along the line you get a steady pattern of highs and lows. This is the "standing wave".

However, in transmission of charge through a wire, the wire is a poor "antennas" and doesn't radiate well. To make a functional antenna, power (i.e., the energy contained in voltages and currents) must be transferred effectively into electromagnetic radiation, where the energy is contained in the electric (E) and magnetic (H) fields [travelling away from an antenna].

A magnetic field, as the result of a moving charge, can also be thought of as the flow of water in a garden hose. As the amount of current flowing increases, the level of magnetic field increases. Magnetic fields are measured in milliGauss (mG). Electric fields are created around appliances and wires wherever a stationary charge, a "voltage", exists. Electric voltage can be thought of as the pressure of water in a garden hose – the higher the voltage, the stronger the electric field strength. Electric field strength is measured in volts per meter (V/m). The strength of an electric field decreases rapidly as you move away from the source. Electric fields can also be shielded by many objects, such as trees or the walls of a building.

Antenna absorbs radio waves and turns them into electrical signals. Antennas are sometimes called receivers. A transmitter an antenna setup that radiates radio waves (i.e., signals; electromagnetic radiation; invisible light). Electron oscillations on the antenna produce electromagnetic radiation in the form of radio waves. To make a good antenna you have to transfer power (the energy is contained in voltages and currents) into electromagnetic radiation (where the energy is contained in the electric field "E" and magnetic field "H") travelling away from the antenna. Antennas can emit radiation and can receive radiation.

The distance from one peak to the next is the wavelength, and the number of peaks passing through a fixed point per unit time is the wave frequency. Electromagnetic radiation is electromagnetic energy in motion. Electrodynamics is the physics of electromagnetic radiation, and electromagnetism is the physical phenomenon associated with the theory of electrodynamics. The electromagnetic field generated from currents and charges (i.e., "sources") is called electromagnetic radiation (EMR), since it radiates from the charges and currents in the source, and has no "feedback" effect on them, and is also not affected directly by them in the present time (rather, it is indirectly produced by a sequences of changes in fields radiating out from them in the past). EMR consists of the radiations in the electromagnetic spectrum, which has been split for government control and commercial application into a series of "bands", including radio waves, microwave, infrared, visible light, ultraviolet light, X-rays, and gamma rays.

24.10 EM radiation and EM waves

The following is a list of notes on EM radiation/waves

1. Radiation is the transfer of energy by way of electromagnetic waves. Waves are what something does, not what something is. Hence, what is waving?
2. Frequency: The frequency of the wave is the number of "crests" (and troughs) [of the wave] that pass a given measurement point within 1 second. In other words, it is the number of complete waves passing a given point in 1 second. And, it has the unit [measurement] of 'Hertz'.
3. Unit: Hertz - 1 wave or cycle, per 1 second, is call a 'hertz'.
4. Energy transfer: The higher the frequency (i.e., the higher the hertz as cycles per 1 second) the higher the amount of energy transferred. Gamma are the shortest (highest) energy "waves" in the current spectrum.
5. Compression and rarefaction: Wavelength is the distance between two consecutive compressions or rarefactions.
6. Note: In general, human vision can detect electromagnetic radiation waves (light) from ~400nm to ~700nm (the visible light region or band of the spectrum).
7. Objects appear to have color because em waves from 400-700nm interact with their molecules. Some wavelengths in the visible spectrum are reflected, and other wavelengths are absorbed. In the case of a green leaf, EM waves from 492-577nm a reflected (which the human eye interprets as green) and the rest of the wavelengths in the visible spectrum are absorbed. Seeing a leaf as green does not give enough information to determine how the leaf reflects UV, microwave, or IR. Everything emits, absorbs, or reflects electromagnetic radiation differently based on its composition. A spectral

signature is a graph showing these interactions across a region of the EM spectrum. Characteristic patterns all for the identification of an object's chemical composition, and determine such physical properties as temperature and density.

8. Sound waves are longitudinal waves - sound travels quickest through a solid.
9. EM waves have a transverse (right angle) and longitudinal (parallel) component.

24.11 Electromagnetic waves

NOTE: *Electromagnetic waves are the geometry taken for the transfer of electromagnetic energy. Mechanical waves (longitudinal for sound and transverse for water) are the geometry taken for the transfer of mechanical energy. A wave is a compression and rarefaction of a medium. It is sometimes said that mechanical waves have a spatial medium (mass), whereas electromagnetic waves have a counterspatial medium (ether).*

Electromagnetic waves (EM Waves) are the oscillating electrical and magnetic fields, acting perpendicular to each other, and propagating through space. EM waves retain their total energy in accordance with the law of conservation of energy. The EM wave spreads out as it travels, which reduces both the field strength and the energy of any section of the EM wave. Total energy of the wave remains the same, however. The relationship between the electrical and magnetic fields at any given point in space is given by Maxwell's equations. An accelerated charge radiates electromagnetic energy in the form of electromagnetic waves. The speed at which energy or signals travel down a cable is actually the speed of the electromagnetic wave, not the movement of electrons. Electromagnetic wave propagation is fast and depends on the dielectric constant of the material. In a vacuum the wave travels at the speed of light and almost that fast in air. An electromagnetic wave is a certain configuration of the electromagnetic field. EM waves carry energy, momentum and angular momentum away from their source particle and can impart those quantities to matter with which they interact. It could also be said that an electromagnetic wave travels through fields and changes them. A field is not the same thing as a wave, but a changing field is experiencing a wave passing through it. And people shortcut this by speaking as if a changing field is a wave. When electric and magnetic fields fluctuate together they lead to formation of the propagating waves called electromagnetic waves. An electromagnetic wave is not constant - it oscillates with time. When an electric (or magnetic) field oscillates, it generates an oscillatory magnetic (electric) field. This oscillatory magnetic (electric) field then generates its own electric (magnetic) field, and back and forth they go until the EM energy in the field is absorbed by matter. This oscillatory electric-magnetic field is an electromagnetic wave. An EM wave can be traveling (e.g. radiation from an antenna) or it can be confined in what is called a standing wave (e.g. the radiation inside a microwave oven). It is the oscillation that makes it a wave. An electromagnetic wave is electromagnetic radiation, is electromagnetic energy in motion, which is described by wave theory. In other words, an EM wave is any EM field that obeys the differential equations governing waves. Technically all EM fields obey this equation, so the definition is usually restricted to fields which have a non-zero frequency component -- that is, fields that oscillate.

An electromagnetic radiation will travel forever, or until it contacts something, in accordance with Newton's first law -- just like any other object in motion.

Electromagnetic waves propagate in vacuum at a maximum speed of 299,792,458 meters per second . For a 12-gauge copper wire carrying a 10-ampere DC current, the speed of electric current (average electron drift velocity) is about 80 centimeters per hour or about 0.0002 meters per second. The speed of electric (electromagnetic) field propagation in copper wire is slower than in vacuum by a factor referred to as the velocity factor. The speed of electromagnetic waves propagate in vacuum is 299,792,458 meters per second. The velocity factor for a 12-gauge copper wire copper wire is about 0.951 (according to this source). Therefore, the speed of electricity in a 12-gauge copper wire is 299,792,458 meters per second x 0.951 or 285,102,627 meters per second. This is about 280,000,000 meters per second which is not very much different from the speed of electromagnetic waves (light) in vacuum.

24.12 Electrical circuits

Electrical circuits provide a means of guiding the transfer of electromagnetic energy (power) via charge carriers in the conductive conduit (i.e., the wire/circuit path).

Electrical circuits in which charges oscillate continuously (alternating currents) will continuously produce both:

1. EM energy through the wire/circuit path, and
2. EM energy that takes the vector path of the magnetic field.

DEFINITIONS: Reactance *is the opposition of a circuit element to a change in current or voltage, due to that element's inductance or capacitance. A built-up electric field resists the change of voltage on the element, while a magnetic field resists the change of current. The electrical* **resistance** *of an electrical conductor is a measure of the difficulty to pass an electric current through that conductor. An ideal resistor has zero reactance, whereas ideal inductors and capacitors have zero resistance – that is, respond to current only by reactance. Note that The magnitude of the reactance of an inductor rises in proportion to a rise in frequency, while* *the magnitude of the reactance of a capacitor decreases in proportion to a rise in frequency (or increases in proportion to wavelength). As frequency goes up, inductive reactance goes up and capacitive reactance goes down.*

24.13 Voltage

A voltage (electromotive force) is required for charges to flow as an electric current. If the voltage difference between two points is zero, there can be no net current between the two points. In other words, charges will flow (as electrical current) through a conductor by applying a voltage across two separated points. The amount of current that flows when voltage is applied is known as amperage.

ANALOGY: *If you have a garden hose and you are trying to push water through it (voltage), you can push as much as you want, but there is a limited amount of flow because the hose is a particular size. Increase the size [of the conduit] and you can increase the amount of flow (amperage).*

Voltage (a.k.a., electric potential difference, electric pressure, electric tension, or electromotive force) is the difference in electric potential energy between two points per unit electric charge. Therein, electrical potential energy is the energy that a charge has when it is at a certain location in an electric field. Each potential difference in a system describes the system's ability to do work. The voltage between two points is equal to the work done per unit of charge against a static electric field to move the test charge between two points and is measured in units of volts (joule per coulomb). A voltage may represent either a source of energy (electromotive force), or lost, used, or stored energy (potential drop). An electromotive force (EMF) is a force that causes electrons (electricity) to flow in a conductor. In a power system, voltage is a measure of the "strength" of an electrical supply.

NOTE: *The higher the voltage, the stronger the electric field.*

Voltage is similar to pressure -- the presence of a potential difference (pressure gradient) drives the electric current, just as the pressure of a pump drives a flow of water. Hence, voltage could be called electrical pressure.

Voltage exists if charges are moving [through] a distance. It is sometimes said that a voltage may exist even when no current is flowing. For example, a disconnected battery has a voltage between its terminals, but because it is disconnected there is no current between the terminals. However, to determine the presence of voltage one must first establish a current; current is required to get/measure voltage. Thus, it is somewhat inaccurate to state that voltage drives the current. From this view, it could be said that a potential difference does not "drive" the current; instead, coulomb force and/or energy is what drives the current. Therein, coulomb force (on a charge) and/or energy can be directly calculated from the potential difference.

NOTE: *Besides superconductors, which can maintain eddy currents flowing in rings with no externally supplied voltage, there can't be currents without voltages, because if there is a current there is a charge moving due to the presence of an electromotive force.*

Unless there is a difference in charge between two points, no field can be established, and hence there is no potential.

The electric charges will gather at the two poles. Positive charges at the cathode and negative charges at the anode. If the two electrodes are not connected by an external conductor they will not be able to leave the surface of the electrodes and they simply accumulate over there producing an open circuit voltage. As soon as the two electrodes are connected by a conductor the charges will flow by the forces of the electric field in the appropriate direction. If the connecting wire has no resistance or almost zero resistance then it will be a short circuit and a huge current will flow only limited by the internal resistance of the battery. If the electrodes are connected by a conductor through a resistance then the current will be limited according to the Ohm's law.

- current (I) = V / (R+r)
- where I is the Current, V is the voltage between the electrodes, R is the external resistance and r is the internal resistance of the battery.

In a battery, the electric field is maintained by the chemical reaction. When connected to a conductor, the charges move through the conductor since it is the path of least resistance.

ANALOGY: *The flow of water through a pipe does depends principally on the pressure difference at the two ends. The flow of charge through a conductor does depends principally upon the charge (pressure) difference at the two ends. It is the pressure (voltage) difference between the two endpoints matters that is of principal significance.*

If electrical work can be done (i.e., there is electrical power), then there is a voltage -- voltage has units of J/C (joules per coulomb). Voltage is expressed and calculated as the difference in electrical energy between two points [in space] per unit electric charge. Voltage is electric [potential] energy per unit charge, measured in joules per coulomb (= volts).

- Voltage (V) = *energy* in joules (J) / *charge* in coulombs (C)
- Voltage (V) = joule (J) / coulomb (C)

- 1V = 1J/C
- Potential = the *ability* to do work.
- Electric potential is the ability to do [electrical] work per electric unit.
- similarly, the electric field is electric force per charge. E = f/q \ f=qE

Notes on voltage:

1. The word "drop" in the term 'voltage drop', comes from the analogy of current being the flow of water and each difference in height that makes the water flow is a drop, a voltage difference. So voltage drop is just a difference in voltage across a component that makes a current flow.
2. A "voltage difference" is the electric potential difference between two points on the circuit, and the current flows in a direction in which the potential difference can be minimized.
3. The second of Kirchhoff's laws tell us that the sum of all the voltages in a circuit must be zero (so, in a simple circuit, the initial voltage from the battery minus all the voltage drops from all the resistors is zero).

Scholarly references

- Calaf, C., Meneveau, C., Meyers, J. (2010). *Large Eddy Simulation study of fully developed wind-turbine array boundary layers*. Phys. Fluids 22, 015110.
- Bassyouni, M., Gutub, S.A. (2013). *Materials selection strategy and surface treatment of polymer composites for wind turbine blades fabrication*. Polymers & Polymer Composites, 21, 463-471.
- Berger L.R., Berger J.A.; Berger (1986). *Countermeasures to Microbiofouling in Simulated Ocean Thermal Energy Conversion Heat Exchangers with Surface and Deep Ocean Waters in Hawaii.* Applied Environmental and Public Health Microbiology. 5 (6): 1186–1198. PMC 239043. PMID 16347076. [aem.asm.org]
- DiPippo, R. (2005). *Geothermal Power Plants: Principles, Applications, Case Studies and Environmental Impact*. Elsevier.
- Griffin, D.A., Ashwill, T.D. (2003). *Alternative Composite Materials for Megawatt-Scale Wind Turbine Blades: Design Considerations and Recommended Testing*. Journal of Solar Energy Engineering. 125 (4): 515. doi:10.1115/1.1629750.
- Mahmoud, S.A., Mohamed, B.S. (2015). *Study on the Performance of Photogalvanic Cell for Solar Energy Conversion and Storage*. Int. J. Electrochem. Sci., 10 (2015) 3340 - 3353. [electrochemsci.org]
- Meyers, J., Meneveau, C. (2012). *Optimal turbine spacing in fully developed wind farm boundary layers*. Wind Energy 15, 305-317 doi:10.1002/we.469
- Rodríguez T. A. et al. (2011). *Wind Turbine Structural Damping Control for Tower Load Reduction*. In: Proulx T. (eds) Civil Engineering Topics, Volume 4. Conference Proceedings of the Society for Experimental Mechanics Series. Springer, New York, NY

Book references

- Londerville, S.B., Baukal, C.E. (Eds.) (2013). *The Coen & Hamworthy Combustion Handbook: Fundamentals for Power, Marine & Industrial Applications*. CRC Press. p12.
- Wheeler, G.J., The design of electronic equipment: a manual for production and manufacturing, Prentice-Hall, 1972

Online references

- Al-Sharif, L.R. (2010). *Mechtronics System Design*. [resources.saylor.org]
- Beaty, W.J. (1996). *What is "Electricity"?* amasci [amasci.com]
- *Conventional and Sustainable Electrical Energy Supply: Overview Characteristics and Comparisons. Electropedia*. Accessed: January 7, 2020. [poweruk.com]
- *Research*. The National Renewable Energy Laboratory. Accessed: January 7, 2020. [nrel.gov]
- *Resources: Technical refence material*. L&S Electronics. Accessed: February 8, 2020. [lselectric.com]
- *TETHYS: Environmental effects of wind and marine renewable energy.* TETHYS. Accessed: January 7, 2020. [tethys.pnnl.gov]
- *The industrial wiki.* Online Dynamic Enterprise Solution for Industry Excellence. Accessed: January 7, 2020. [myodesie.com]
- Lower, S. (2016). *Understanding Entropy*. Chem 1 Virtual Textbook. [chem1.com]
- *What is the difference between an induction motor and a synchronous motor?* Quora. Accessed: February 8, 2020. [quora.com]
- *What is the difference between torque and moment?* Quora. Accessed: February 8, 2020. [quora.com]

TABLES

Table 13. **Life Support > Power > Primary**: *Primary energy "generating" sources accompanied by a description of where the energy is derived from.*

Primary Energy Generating Sources			
Energy source generators	**Energy from / transport by**	**Conversion process**	**Power depends on** (Conversion rate of generator and ...)
Geothermal (thermal)	Heat from inside the earth	Turbine	Temperature
Wind (atmospheric current)	Atmospheric currents derived from the earth's rotation and exposure to radiant energy from the sun	Turbine	Wind speed ~v^3
Wave (water + wind)	Wind waves in a body of water	Wave device / Turbine	Wave height (H^2) and wave period
Tidal (water + solar system gravity/electrostatic force)	Daily cyclical movement of a body of water	Tidal device / Turbine	Height squared (H^2) and flow speed (cubic)
Hydro (water + planetary gravity/electrostatic force)	Uni-directional flowing movement of water	Dam / Turbine	Height squared (H^2) and flow speed (cubic)
Hydrolysis (water)	Chemical reaction of an organic molecule breaking down in water	Reactor	Delocalization across the C9-N15 bond and steric effects
Solar (radiation)	Solar non-thermal radiation from the sun	Photovoltaic	~ annual radiation
Solar (thermal)	Solar thermal radiation from the sun	Collector	~ annual radiation
Animate (animal)	Animal movement	Animal movement	Species, sex, and strength/health of organism
Coal (solid hydrocarbon)	Combustion of organic rocks	Combustion	Heat content of the fuel
Oil (liquid hydrocarbon)	Combustion of organic liquids	Combustion	Heat content of the fuel
Gas (gas hydrocarbon; a.k.a., "natural gas")	Combustion of organic gases	Combustion	Heat content of the fuel
Biomass (plants & animal waste)	Combustion of plant-derived materials	Combustion	Heat content of the fuel
Nuclear (atomic)	Heat from fission of large atoms or fusion of small atoms	Reactor	Temperature

Table 14. **Life Support > Power > Energy Conversion**: *Example conversions with efficiency notation.*

Energy Conversions			
Converter	**Form Of Input Energy**	**Form Of Output Energy**	**Efficiency**
petrol engine	chemical	mechanical	η
diesel engine	chemical	mechanical	η
electric motor	electrical	mechanical	η
boiler & turbine	thermal	mechanical	η
hydraulic pump	mechanical	potential	η
hydro turbine	potential	mechanical	η
hydro turbine	mechanical	electrical	η
generator	mechanical	electrical	η
battery	chemical	electrical	η
solar cell	radiation	electrical	η
solar collector	radiation	thermal	η
electric lamp	electrical	light	η
waterpump	mechanical	potential	η
water heater	electrical	thermal	η
gas stove	chemical	thermal	η

TABLES

Table 15. <u>Life Support > Power > Energy Type</u> *Elaborated list of energy forms and energy types with accompanying descriptions. Note that wave energies (such as radiant or sound energy), kinetic energy, and rest energy are each greater than or equal to zero because they are measured in comparison to a base state of zero energy: "no wave", "no motion", and "no inertia", respectively.*

Types of energy	Description of energy Type
Kinetic	(≥0), that of the motion of a body.
Potential	(≥0), that of the position of a body relative to the zero plane of inertia of that body.
Forms of Energy	Description of energy Form
Mechanical	The sum of (usually macroscopic) kinetic and potential energies. The energy of motion (every moving object). Usually visible.
> *Mechanical [wave]*	(≥0), a form of mechanical energy propagated by a material's oscillations -- 'acoustic energy' is called sound.
Chemical	That contained in molecules (molecular bonds).
Electric	That from electric fields.
Magnetic	That from magnetic fields.
Electromagnetic	(≥0), that of electromagnetic radiation including light -- 'optical energy' is called light; 'radiant energy' carried by light.
Nuclear	That of binding nucleons to form the atomic nucleus.
> *Ionization*	That of binding an electron to its atom or molecule.
Thermal	A microscopic, disordered equivalent of mechanical energy. Expressed as heat.
> *Heat*	(≥0), the microscopic motion of molecules. An amount of thermal energy being transferred (in a given process) in the direction of decreasing temperature -- heat is a form of energy; temperature is a measurement of heat.
Forms of power *(Thermodynamically, only 2 forms of power)*	Description of power type
Mechanical power (work[ing])	The rate at which "work" is done. Mechanical energy used per unit time.
Thermal power (heat[ing])	The rate at energy is transferred via heat. Thermal power is the measure of thermal energy used per unit time. It is the rate of heat transfer or heat flow rate.

Table 16. <u>Life Support > Power > Energy Kinetic</u>*: Forms of kinetic energy (classified by type of motion).*

Kinetic Energy Forms	Motion	Examples And Subtypes Of This Form Of Energy
Mechanical [motion]	motion of macroscopic objects/substances;	machines, muscles, projectiles, wind, flowing water, mechanical waves, sound (acoustic, longitudinal waves), ...
Thermal [motion]	vibratory motion (vibration) of microscopic particles of matter (molecules, atoms, ions) --	heat, fire, geothermal, ...
Electrical [motion]	flow of charges (electrons, protons, ions)	electric current, AC and DC circuits, ...
Electromagnetic [motion]	disturbance propagating through electric and magnetic fields or the motion of photons	the electromagnetic spectrum [banded into radio waves, microwaves, x-rays, ...]

Table 17. <u>Life Support > Power > Energy Potential</u>*: Forms of potential energy (classified by type of mathematical field).*

Potential energy forms "Forces"	Quantity in field	Examples and subtypes of this form of energy
Gravitational [force field]	mass	roller coaster, waterwheel, hydroelectric reservoir, ...
Electromagnetic [force field]	charge	electric, magnetic, chemical, elastic, ...
Strong nuclear [force field]	color charge	nuclear reactors, nuclear weapons, ...
Weak nuclear [force field]	lepton number	radioactive decay, mass change, ...
Chemical [force field]	ion[ic charge] - atoms and molecules separated into ions (ionic bonds)	endothermic and exothermic reactions

TABLES

Table 18. **Life Support > Power > Energy Flow**: *Energy flow breakdown examples.*

Energy Flow Breakdown Examples		
Energy Stage	**Technology Used**	**Example of objects in stage of flow**
Primary	-	coal, wood, hydro, dung, oil, etc
-	Conversion	power plant, kiln, refinery, digester
Secondary	-	refined oil, electricity, biogas
-	Transport/Transmission	carriage, pipes, wires
Final	-	diesel oil, charcoal, electricity, biogas
-	Conversion	motors, heaters, stoves
Useful	-	heat, shaft power

Table 20. **Life Support > Power > Energy Transformation**: *Energy transformation: coal fired power plant example.*

Coal-Fired Power Plant Example Of Energy Transformations	
Energy Transformation	**Description Of Transformation**
Chemical energy	coal converted to thermal energy in the exhaust gases of combustion
Thermal energy	the exhaust gases converted into thermal energy of steam through the heat exchanger
Thermal energy	steam converted to mechanical energy in the turbine
Mechanical energy	turbine motion converted to electrical energy by the generator, which is the ultimate output
In such a system, the first and fourth step are highly efficient, but the second and third steps are less efficient. The most efficient gas-fired electrical power stations can achieve 50% conversion efficiency. Oil- and coal-fired stations achieve less.	

Table 19. **Life Support > Power > Energy Transformation**: *Energy transformation types and descriptions.*

Energy Type	Description of Energy Transformation
Thermoelectric	Heat > electric energy
Geothermal power	Heat > electric energy
Heat engine	Heat > mechanical energy
Ocean thermal power	Heat > electric energy
Hydroelectric dams	Gravitational potential energy > electric energy
Electric generator	Kinetic energy or mechanical work > electric energy
Fuel cells	Chemical energy > electric energy
Battery	Chemical energy > electric energy
Fire	Chemical energy > heat and light
Wave power	Mechanical energy > electric energy
Wind power	Mechanical energy > electric or mechanical energy
Piezoelectrics	Mechanical ("strain") energy > electric energy
Acoustoelectrics	Mechanical ("acoustic/sound") energy > electric energy
Friction	Kinetic energy > heat
Heater	Electric energy > heat

TABLES

Table 21. **Life Support > Power > Physics**: *Physics > Electrostatics > Charges. Opposite charges attract. When there is an equal # of opposite charges there is "balance", giving the atomic system an overall neutral (zero) charge.*

Name	**Signifier**	**Unit**	**Relationship**
Proton	+ (positive)	charge	
Electron	- (negative)	charge	
Name	**Signifier**	**Unit**	**Relationship**
Positive charge		Cations	Possesses more protons than electrons; higher electric potential
Negative charge		Anions	Possesses more electrons than protons; lower electric potential
Uncharged / electrically neutral		Neutrino	Equal numbers of protons and electrons; no net electrical charge; equipotential throughout

Table 22. **Life Support > Power > Physics Energy**: *This table depicts the different conceptualizations of energy, the incorrect and correct scientific conceptions, and their information analogues.*

Conception	Incorrect idea	Scientific concept	Information Analogue
Energy as agent	Energy causes things to happen. It makes an action happen and can be stored inside a physical thing.	Energy does not cause events to happen. However, when events happen, there is always a transfer of energy between interacting physical things.	Information does not cause events to happen. However, when events happen, there is always a transfer of information between interacting physical things.
Energy as action	Energy is an action or activity, like burning, bubbling, running, and bouncing.	Actions are visible experiences that energy is transferring.	Actions are visible experiences that information is transferring.
Energy as form	Energy has multiple forms depending upon its location in the physical world.	Energy does not have different forms or location in the physical world, but there are different types (or modes) of energy transfer - energy is transferred within and between "carriers" in different ways.	Information does not have different forms or location in the physical world, but there are different types (or modes) of information transfer - information is transferred within and between "carriers" in different ways.

Table 23. **Life Support > Power > Prime**: *Primer movers as types of work and power.*

Prime mover	Type of work	Type of power
Force	Mechanical work	Mechanical power
Pressure	Fluid work	Fluid power
Voltage	Electrical work	Electric power

TABLES

Table 24. **Life Support > Power > Types**: *Power types and their properties.*

	Power Type			
Property	**Electrical**	**Mechanical**	**Pneumatic**	**Hydraulic**
Energy transition	Turbine	IC engines, electrical energy is used to drive motors	Electrical energy is used to drive compressors and other equipment	IC engines, electric motor, air turbine used to drive hydraulic pump
Medium	Energy is transferred wirelessly or wired	Energy is transferred through levers, gears, and shafts	Compressed air/gas in pipes and hoses	Pressurized liquid in pipes and hoses
Energy storage	Batteries	Variable frequency drives	Reservoir, air tank, pneumatic valves	Accumulators, hydraulic valves
Transmitters	Wireless and wired transmitters	Transmitted through mechanical components like levers, gears, cams, screws, etc.	Transmitted through pneumatic cylinders, rotary devices, and rotary actuators	Transmitted through hydralic
Leakage	Stray voltage and ground currents	N/A	Contamination relative to gas used	Contamination relative to liquid used
Energy transmission/ distribution	Unlimited with power loss	Short distance	Up to 100m flow rate v = 2-6 m/s signal speed up to 1000 m/s	Up to 1000m flow rate v = 20-40 m/s signal speed 20-40 m/s
Operating speed			v = 0.5 m/s	v = 1.5 m/s
Power supply input	Low	Variable	High	Very high

Table 25. **Life Support > Power > Circuit/Ground**: *Grounding system comparison table. In the 1999 Edition of the NEC, impedance grounded systems were considered to be ungrounded systems.*

Grounded Systems				
NEC reference	**Required grounded systems**	**Pros**	**Cons**	**Use**
Article 250.20(A)	(Solidly grounded) ac systems less than 50V	Greater safety Prevents insulation damage from over-voltages from line-to-ground faults during resonant ground faults Easy detection; faster mean time to repair	Higher fault levels; validate circuit breakers rated for bolted fault	Equipment such as window shades, BAS, and some fire alarm systems
Article 250.20(B)	(Solidly grounded) ac systems 50V to 1,000V			Residential single-phase 120V/240V Commercial and light industrial facilities with 3-phase, 480V:208 Y/120Vsystems
Article 250.20(C)	(Solidly grounded) ac systems 1,000V and above			Medium-voltage transmission lines
Article 250.20(D)	Impedence grounded systems	Provides operational continuity during a line-to-ground fault; equipment doesn't shut down Introduction of the resistance can control the higher fault levels present in the solidly grounded system	Must be engineered to match facility capacitance	Industrial plants; mills Large data centers Medium-voltage cables
Ungrounded Systems				
NEC reference	**Allowed ungrounded systems**	**Pros**	**Cons**	**Use**
Article 250.21(A)	General: systems deemed to be a higher safety risk to automatic shut down, suich as blast furnaces	Provides operational continuity during a line-to-ground fault. Equipment doesn't shut down Cheaper to install	Primary line-to-ground transietns are passed through transfomers unattenuated	Steel manufacturing Industrial plants Pulp and paper

TABLES

Table 26. Life Support > Power > Mechanical Electric*: Difference Between Induction and Synchronous motors and generators is explained with the help of various factors.*

Basis Of Difference	Synchronous Motor	Induction Motor (Asynchronous Motor)	Synchronous Generator (Alternator)	Induction Generator (Asynchronous Generator)
Type of excitation	A synchronous motor is a doubly excited machine.	An induction motor is a single excited machine.	-	-
Frequency	-	-	Frequency is determined by the rotational speed of the generator's shaft -- faster rotation of the shaft generates a higher frequency.	-
Supply system	Its armature winding is energized from an AC source and its field winding from a DC source.	Its stator winding is energized from an AC source.	-	-
Speed	It always runs at synchronous speed. The speed is independent of load. Synchronous motors are used where constant running speed is the governing factor. In a synchronous motor the rotor and magnetic field rotate at the same speed.	If the load increases, the speed of the induction motor decreases. It is always less than the synchronous speed. The operation of the induction motor depends on relative motion as the difference in speed between the rotor and the rotating magnetic field. This relative motion induces and EMF in the rotor.	-	-
Starting	It is not self starting. It has to be run up to synchronous speed by any means before it can be synchronized to AC supply.	Induction motor has self starting torque.	-	Usually not started without an energized connection to the electric power grid, unless they are designed to work with a battery bank energy storage system.
Operation	A synchronous motor can be operated with lagging and leading power by changing its excitation.	An induction motor operates only at a lagging power factor. At high loads the power factor becomes very poor.	-	-
Usage	It can be used for power factor correction in addition to supplying torque to drive mechanical loads.	An induction motor is used for driving mechanical loads only.	-	-
Efficiency	It is more efficient than an induction motor of the same output and voltage rating.	Its efficiency is lesser than that of the synchronous motor of the same output and the voltage rating.	-	-

Table 27. Life Support > Power > Solar Electric*: Direct solar to electric conversion types.*

No	Types	Characteristics
1	Photoemissive	Light interacting with a cathode causes electrons to be emitted from the cathode surface.
2	Photoconductive	The resistance of a material is changed when it is illuminated.
3	Photovoltaic	Light interacting with the junction of two exposed substances generates an output voltage proportional to light intensity.
4	Photomagnetic	Light interacting with a dynamic magnetic field causes a voltage.
5	Photogalvanic	Light interacting with a material produces a chemical action that causes voltage.
6	Photoelectrochemical	Light interacting with a material produces a chemical action that causes voltage.
7	Bio-photoelectrochemical	Light interacting with an organic material produces a chemical action that causes a voltage.

TABLES

Table 28. Life Support > Power > Storage: *Typical values of specific energy and energy density.*

Common storage materials (fuels)				
Energy source	**Energy reaction type (to release)**	**Density Kg/m3**	**Specific energy mj/kg**	**Energy density Mj/m3**
Coal (anthracite)	Chemical	1350	-27	-36,450
Coal (lignite)	Chemical	801	-15	-12,015
Wood	Chemical	600	-15	-9,000
Common storage hydrocarbons (fuel alkanes)				
Hydrocarbon (alkane) storage	**Energy reaction type (to release)**	**Density Kg/m3**	**Specific energy mj/kg**	**Energy density Mj/m3**
Methane (ch4)	Chemical	423	-55.5	-23,529
Ethane (c2h6)	Chemical	545	-51.8	-28,246
Propane (c3h8)	Chemical	585	-50.3	-29,449
Butane (c4h10)	Chemical	601	-49.5	-29,729
Pentane (c5h12)	Chemical	621	-48.7	-30,223
Hexane (c6h14)	Chemical	655	-48.3	-31,633
Heptane (c7h16)	Chemical	680	-48.1	-32,690
Octane (c8h16)	Chemical	698	-47.9	-33,433
Decane (c10h22; kerosene)				
Common storage alcohols (fuels)				
Alcohol storage	**Energy reaction type (to release)**	**Density Kg/m3**	**Specific energy Mj/kg**	**Energy density Mj/m3**
Methanol (ch3oh)	Chemical	787	-22.7	-17,855
Ethanol (ch3ch2oh)	Chemical	785	-29.7	-23,278
1-Propanol (ch3(ch2)2oh)	Chemical	800	-33.6	-26,902
Common storage devices (battery-like devices)				
Device storage	**Energy reaction type (to release)**	**Density Kg/m3**	**Specific energy Mj/kg**	**Energy density Mj/m3**
Lithium battery (non-rechargeable)	Electrochemical		1.8	4.32
Lithium-ion battery (rechargeable)	Electrochemical		0.36–0.875	0.9-2.63
Alkaline battery	Electrochemical		0.5	1.3
Lead-acid battery	Electrochemical		0.17	0.56
Nickel-metal hydride battery	Electrochemical		0.288	0.504-1.08
Supercapacitor (EDLC)	Electrical (electrostatic)		0.01-0.036	0.06-0.05

TABLES

Table 29. Life Support > Power > Storage: *Functional differences between a battery and capacitor.*

Function	Capacitor	Battery
Charge time	1-10 sec	10-60min
Cycle life	1 million or 30,000 h	500 and higher
Cell voltage	2.3 to 2.75V	3.6-3.7V
Specific energy (Wh/kg)	5 (typical)	100-200
Specific power (W/kg)	Up to 10,000	1,000-3,000
Service life (in place) in years	10-15 years	5-10 years
Charge temperature (between two values)	-40 to 65˚C	0 to 45˚C
Discharge temperature (between two values)	-40 to 65˚C	-20 to 60˚C

Table 30. Life Support > Power > Storage: *Overview of sensible, latent, and thermochemical processes using salt.*

Temperature level	Salt type	Test type
<0 ˚C	Water-salt mixtures	PCM slurry
0-100 ˚C	Melting of salt hydrates in crystallization water	PCM
40-300 ˚C	Dehydration of salt hydrates	TCS
40-150 ˚C	Absorption in concentrated salt solutions	TCS
120-500 ˚C	Solid–liquid conversion in anhydrous salts	PCM
100-800 ˚C	Anhydrous molten salts	Sensible
100-800 ˚C	Anhydrous solid salts	Sensible
100-800 ˚C	Solid–solid conversion in anhydrous salts	PCM

Table 31. Life Support > Power > Storage: *Electrochemical capacitor types.*

Electrochemical capacitor type	Symmetric aqueous	Symmetric organic	Asymmetric aqueous	Asymmetric organic
Energy density				
Power performance				
Self discharge rate				
Low-temp discharge				
Packaging				
Voltage balance				
Cell voltage				
Operating temperature limits				

Table 32. Life Support > Power > Conversion: *Energy "transformation".*

Energy Resources	>>>	**Technological Equipment**	>>>	**Usable Energy**
For example, moving water, biomass, wind, sunshine, the Earth)		For example, hydroelectric and wind turbines, stoves and furnaces, photovoltaic panels)		For example, electricity, steam, heat, biofuels

TABLES

Table 33. Life Support > Power > Load*: Energy requirements of a device/load.*

Device/load (examples)	Instantaneous power requirement/ demand/usage "rating" P=E/t	Energy requirement/usage:		
		If run for 1 second (energy consumed)	If run for 1 minute (energy consumed)	If run for 1 hour (energy consumed)
60W light bulb	60w 60J/s	60w 60J/s		
Laptop	40-80w 40J/s - 80J/s	0-80w 40J/s - 80J/s		

Table 34. Life Support > Power > Load*: Example electrical energy demand profile.*

Name and amount of device	x	Device's required POWER (rate of demand in 1 minute in Watts)	x	TIME the device is used (in hours)	=	Total energy consumption of device (in watt-hours)	Power usage in time
1 Light bulb	x	100 Watts (1x100 = 100W or .1kW)	x	10 hours	=	1,000 Watt-hours or 1kWh	1kW used in 10 hours
10 Light bulbs	x	100 Watts (10x100 = 1000W or 1kW)	x	1 hour	=	1,000 Watt-hours or 1kWh	1kW used in 1 hour; 10 times more demand than 1 light bulb over 10 hours

Table 35. Life Support > Power > Conversion Electric*: Electric power conversion classified according to whether the input and output are alternating current (AC) or direct current (DC). A power converter is an electrical or electro-mechanical device for converting electrical energy.*

Electric power conversion			
DC to DC	DC to AC	AC to DC	AC to AC
Dc-to-DC converter	Inverter	Rectifier	Transformer / autotransformer
Voltage regulator		Mains power supply unit (PSU)	Voltage converter
Linear regulator		Switched-mode power supply	Voltage regulator
			Cycloconverter
			Variable-frequency transformer

Table 36. Life Support > Power > Electricity*: AC and DC device differences.*

AC Devices (~1950)	DC Devices (~2000)
Electric typewriters and adding machines	Computing and printing equipment
Teleprinter	Telecommunication systems
Early fluorescent lighting	Advanced fluorescent lighting with electronic ballast, gas discharge lighting, LEDs
Radios, early televisions	HDTV's, CD Players, smartphones
Record players	CD players and game consoles
Electric ovens	Microwave ovens
Fans and furnaces	Electronically controlled HVAC

Life Support: Medical Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: medical, health service, first response, first aid, lifeform restoration

Abstract

Human life, like all biological life, is composed of consciousness animating a biological life [organism] form. In this dimension of reality, given the technologies available, the human biological form can be injured and will not become whole again instantly [and through mental thought alone]. Injury and dis-ease can degrade the quality of life of an individual at the same time degrading the quality of life of those in interrelationship with the injured and/or intensely suffering individual. Even in a community-type society, it is still possible for social conflict to emerge such that medical services may need to include individuals trained to stop violence. A community-type society perceives violence as an accident in society to be individually recovered from and socially restored to a dynamic where the violence is unlikely to present itself in the future. Violence is a medical issue, as are physical injuries, skills and knowledge are similar in some regards and different in others. Here, there are no police-type roles as seen in the market-State. Issues that cause harm and trigger life-oriented restorative processes and protocols are likely to be handled different depending on what values (on a values circumplex) are being actively encoded within that society. Some values are more likely to engage punitive-type behaviors; whereas other values are more likely to engage restorative-type behaviors. A medical service is, by definition, of the restorative type. A medical service provides the capability of dealing with medical problems. A medical service is, in part, a disaster recovery plan for life forms. A medical system involves medical knowledge, medical procedures, and medical drugs, but healthy well-being is primarily achieved by facilitating individual happiness with respect to whatever might have distressed them into illness. Thus, good health-care places enormous emphasis on dealing with the cause of illness, as opposed to remedial treatment of the symptoms. Humans can be injured physically and mentally, together.

Graphical Abstract

Not Currently Applicable

1 Introduction

In its scientific context, medicine is the science of creating and restoring healthy functioning and facilitation of biological resilience. Pain means something is wrong; pain has a function. Pain is a signal of something, often, tension. The more a source of pain continues to generate the pain, the more disease there will be. There are emergency cases of pain, such as a vehicular accident where humans are physically injured, and there are cases of pain, such as overeating and arthritis. Some medical problems are life-threatening, and others are not. Because there are life-threatening medical problems, there is the necessity for a medical service to be continuously operative and available to respond. Some medical procedures can be carried out remotely, but others require structures and technologies positioned at fixed locations.

A proper medical system restores the individual and the social to wholeness, together. Science and ancestral wisdom can be used together to address the whole personal, and their experience throughout time.

Human performance can be compromised by many aspects of a task, including:

1. Body posture
2. Physical and biomedical composition
3. Sleep restoration (awakeness)

1.1 *Disease and the environment*

Dis-ease can come from intentional re-arrangement of the environment. For instance, there is no moderation when humans are surrounded by, and consume, hyper palatable, which they are wired to eat more of. It is basic biological wiring to eat the available and to overeat on highly palatable foods (especially hyper-palatable food that is also nutrient poor). In general someone where to make their meals enjoyable, full of real food, but not over the top palatable (i.e., to the point that overeating becomes more likely). In place of this type of food, some societal arrangement make food with hyper-palatability and low nutrients more accessible, thus leading to behaviors that generate dis-ease states within the environment, because real signals are being obstructed. Humans are designed to crave substances that are good for them (i.e., that the body needs). But when food service systems are designed outside of nature, then the bodies craving mechanism (taste) can quickly become misinformed and aberrant such that people crave foods and food-like substance that do not satisfactorily meet their bodies needs.

1.2 *Medical technology inventory*

Medical technology and inventory
Team coordination/control protocols

1. Emergency medical response
   A. Pharmaceuticals
   B. Decision support (for humans)
   C. Semi-automated systems
   D. Fully-automated systems
   E. Medical studies
   F. Operations
2. Equipment
   A. Laboratory diagnostic equipment
      1. Clinical chemistry
      2. Hematology
      3. Pathology
      4. Microbiology
      5. Hematology, and endocrinology
   B. Imaging diagnostic equipment
      1. Radiographic
      2. Magnetic resonance
      3. Ultrasound
   C. Minimally invasive or non-invasive monitors
      1. Electrocardiograph, blood pressure, oxygen saturation, etc.
      2. Equipment and protocols to provide rescue, resuscitation, stabilization, and transport.
   D. Surgery
      1. Microsurgery/micro-therapeutics equipment and protocols.
      2. Specialized surgery
   E. Fluid therapy systems including infusion pumps, on-site production of fluids, nutritional support, blood, and blood component replacement.
   F. Methods for biomonitoring
   G. Medical waste management
   H. Medical storage systems for samples, pharmaceuticals, and other perishable items.
   I. Medical energy management

1.3 *Medical response*

INSIGHT: *People are more likely to heal faster when they have an aesthetic view from their window as opposed to a hospital bed with a view over the parking lot.*

1.3.1 The first responders

A.k.a., The first responding intersystem team.

Must account for proximity location.

Life Support: Cultivation Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: cultivation, biological cultivation, food cultivation, food service, food production, biological nutrition, materials cultivation, biological materials production

Abstract

Human life, like all biological life, is composed of prior animated life and minerals. Food is organisms (i.e., their bodies), including animal, vegetable, fungal, and bacterial, or it is things that come out of the bodies of these organisms. In nature everything is food for something else. It is possible to produce a service that cultivates food and other useful materials from organisms (e.g., textiles). Here, the mimicking of natural patterns is often the optimal strategy. Technology can be used to provide means and timeframes for cultivation that are impossible without the technology (e.g., hydroponics). Food has flavor. Food can, or cannot, be appropriate in nutrient content for the consuming organism. Textile and food products are the output of the cultivation service. Some foods are processed prior to distribution and others are not. In a community-type society, individuals can prepare and cook their own food, or they can rely on automated services. Similarly, material cultivation system may rely on human effort (e.g., such as personal gardens), others may involve a combination of human team and machine interaction, and even others may be fully automated operations. In general, the cultivation service involves the cultivation of resources from other organisms. There is a second dimension to the cultivation service, mining. Mining isn't really a form of cultivation because it doesn't involve the cultivation of living organisms. However, it does involve the "cultivation" of minerals (including hydrocarbons) from the earth. Hence, the cultivation service involves both the cultivation of other organisms and the extraction (mining) of compounds from the earth. Mining may include the collection of matter in its various forms (e.g., gaseous, liquid, and solid).

Graphical Abstract

Not Currently Applicable

1 Life cultivation

The fundamental difference is in the way animals and plants take in carbon to form organic compounds. Plants are autotrophs, which means that they meet their carbon requirements solely from carbon dioxide in the atmosphere, or from water in the case of water-dwelling plants. Animals, being heterotrophs, are unable to make their own organic molecules and so must take them in ready-made by eating plants and other animals.

1.1 Plant life requirements

The nine cardinal parameters for growing plants:

1. Foliar zone environment parameters:
   A. Light
   B. Temperature
   C. Humidity
   D. Wind (air velocity)
   E. CO_2
2. Root zone environment parameters:
   A. Temperature (root-zone)
   B. Water
   C. Oxygen
   D. Nutrients

1.1.1 Permaculture

Every action taken in a permacultural system is taken with thought to other systems, and the action often carries with it a, secondary useful purpose. The garden produces food, but it also produces medicines and stabilizes the ecology, produces smells, and produce nutrients for the local ecology.
Regenerative farming practices may involve, but are not limited to:

1. Rotational grazing.
2. Intensive silvo pasture - silvo meaning tree, and silvo pasture meaning trees in the pasture. Planting trees in the pasture. Young trees need to be protected because otherwise grazing animals will kill them by eating them or rubbing/scratching on them. The right species. Will protect the animals and provide shade. If it is the right species it increases the productivity of the pasture. There are many types of silvo pasture. In particular, there is the intensive silvo pasture - planting trees every square meter or closer, and interplanted with pasture grasses. Silvo pastures require a hydrated pasture and wont work on land with very little water.
3. Pastures planted with pasture grasses. The animals enter a pasture and graze them down to the found whereupon they are moved into another paddock. The grass repeats and you bring the animals in again.
4. Carbon farming are a group of methods that remove excess carbon dioxide and store it in the form of perennial plants (like woody vegetation) and organic matter in the soil.
5. Agroforestry - farming that incorporates trees and/or woody plants.

1.1.1.1 Permacultural terrain reconstruction

Permacultural terrain reconstruction involves actions that recover a landscape to become more capable of time of sustaining and producing life:

1. Plants take root and bamboo rots away.
2. Circular, ecological cultivation systems
   A. Regenerative farming, circular symbiotic farming, etc.
3. System sub-classifications
   A. Closed-loop systems (e.g., closed hydroponic)
   B. Open-loop systems (e.g., outdoor garden)
4. Permacultural habitats
   A. Gardens – vegetable organisms
   B. Orchards – vegetable organisms
   C. Greenhouses – vegetable organisms
   D. Aquaculture – aquatic life
   E. Fungi – Fungi organisms as fruiting bodies and decomposers also
   F. Insects – insect farms

1.1.1.2 Killing organisms

The ways of killing organisms include, but are not limited to:

1. Chemical "-icides"
   A. Pesticide
   B. Fungicide
   C. Insecticide
   D. Larvicide
   E. Herbicide
   F. Rodenticide
2. Mechanical (e.g., crushing, gunshot, etc.)

1.1.2 Plant mono-agriculture

Growing acres of mono-agriculture will deny that land to the animals that would otherwise have lived, and eaten, and thrived there.

1.2 Animal life requirements

The seven cardinal parameters for growing animals:

1. Boundary transference environment:
   A. Light
   B. Temperature
   C. Humidity

D. Wind (air velocity)
E. Atmospheric composition with at least O_2
F. Water
G. Nutrients

1.3 Animal stewardship and regenerative/ cyclical farming

Animal stewardship involves the caretaking and use of other animals. Regenerative and cyclical farming (a.k.a., symbiotic farming, etc.) necessitates the caretaking of other animals and the total ecology.

Through design, cities may become unique ecosystems that facilitate the caretaking of the total planetary ecological system. By highly controlling and coordinating the flow of resources within the city (habitat service system), it is possible to utilize the principles of natural systems in to cycle and regenerate the continuous flow of resources upon the planet for global human fulfillment.

1.4 Animal ecological functions

Goats can be "run" through nature surrounding the habitat service system as a fire mitigation strategy. They help eliminate excessive fuel in the area by eating invasive plant species. Can be used to eliminate noxious weeds. This will mitigate the risk of wildfire by reducing the amount of readily available flash fuel. Put an electric fence around them for each temporary grazing area. The alternative to prevent growth is weeding, mowing, or chemical usage. Depending upon the area, they can clear about 2 acres a day. May want to keep guard dogs around to protect them from predators.

Pigs can recycle some of the unused production from the farm/cultivation system. A certain number of the eggs from the hens won't pass inspection, and so feed them to the pigs.

Pigs are forest animals, as opposed to pasture animals like cows and sheep. They do a great job of clearing out the scrubb so that it isn't a jungle right beyond the pasture land. Goats can go in the forest with the hogs. Hogs will eat the goat faeces. When they do that, they will pick up some nutrients. Heterosis. The pigs eating the faeces of the goats breaks the lifecycle of an internal parasite called barberpole worm. The eggs are passed through the goat, its faeces, it pupates in the earth, then the goats graze the grass and reinfect themselves. When the pigs eat the faeces they break the lifecycle.

2 Food

Any biological organism exists as long as a stream of energy passes through it in the form of food for animals or electromagnetic radiation in combination with inorganic compounds for plants. Food is lifeforms or things secreted by lifeforms. When humans eat food they eat the tissues of living, or recently living, things. Individuals might eat the colonies of the bodies of bacteria when we eat fermented food. Humans eat life and life eats life, and there is no present way around that.

When an organism eats food it is eating other organisms (their bodies), or it is eating things that come out of other organisms, as well as raw physical elements. The human organism eats living things or formerly living things, necessarily, to remain alive. Material bodies consume and make use of material substances, and when the amounts are off or the substances are not right, then the conscious organism doesn't feel its best.

In a closed ecosystem nothing goes to waste; not even waste (e.g., excrement). Excrement is food for other organisms.

There are sub-kingdoms of organism:

- 4 Kingdoms: animals, plants, fungi, and bacteria.
- 6 Kingdoms (biological sciences): archaebacteria, eubacteria, fungi, protista, plants and animals.

> **INSIGHT:** *Places where food is present may determine the flow of persons over a landscape. Food could be seen as 'place', and connection to 'place' - eating the landscape you live in*

2.1 Cultivation for food

Humans are an omnivore and through various technologies they are able to render a lot of foods that would otherwise be inedible on the wild landscape, edible. Food processing technologies include both mechanical technologies and chemical technologies. Most wild food has to be processed for someone prior to consumption. Berries and other fruits not so much, but for most staple foods there is a process that must be gone through to render them edible. And, that is one of the unique things about the human animal. Humans can take foods that are otherwise inedible to them, and through technologies they have developed (e.g., mechanical or chemical technologies), they are able to render a lot of material (food) edible, and in some cases, more nutritious. Hence, humanity is not just an omnivore, but a technological omnivore. Not highly complex technologies like smart phones and 3d printers, but basic technologies like grinding and cooking that humanity has had for millennia. These technologies allow humanity to prepare and consume foods that might otherwise be too tough, unpalatable, or have chemicals in them that need to be deactivated through heat for them to be safely digested.

The following are kingdoms/types of life that humans can eat:

1. Insects (entamophagy)
2. Fish and other marine animals
3. Land animals (meat comes from wild animals and from animal stewardship; meat consumption is a 200000 year practice)
4. Plants
5. Fungi
6. Bacteria

2.1.1 Cultivation separation

Someone can't (or at least, shouldn't) harvest a plant like dandelion greens where people are walking dogs. So, there are areas where dogs cannot go; there is separation of objects through structure where necessary in order to ensure cultivation plans are successful. That which cultivated foods and other resources need to be separated from may include wild animals, it may include animals within and around humans, and it may include separation from humans (e.g., so as not to contaminate a food or food process).

2.1.2 Food, flavor, and diet

> **APHORISM:** *If you don't like what comes out - change what goes in.*

Humans are designed to crave the foods that bring them what they need in terms of micro and macro-nutrient profiles. And, flavor is the mechanism. And, the flavor is what has changed. Since the 1950s,the flavor of whole foods has been getting progressively blander and blander. At the same time, the flavor of the foods that are not good for Hans, the processed foods has been getting more delicious and more irresistible. When flavor is changed, so is the incentive to eat. It is no surprise in the early 21st century that people are eating the wrong foods. It is a perfect storm of making the healthy whole foods taste bland, and the unhealthy foods take delicious and taste such that many people just can't stop eating it. Often, if it wasn't for the added flavoring chemicals people wouldn't eat the substrate.

There is "food" that has the appearance (flavor or look) of food, but not the substance of it.

2.2 Hygiene and food cultivation

Hygiene is a concern with food residue, bacteria and other potential disease causing organisms.

2.3 Food storage

Food storage and preservation techniques include, but are not limited to:

1. Gas introduction
   A. Ozone
   B. Nitrogen
2. Vacuum
   A. Sterilization of vacuumed containers (e.g., cans)
3. Air depressor locks
4. Fermentation
5. Dehydration (desiccation)
   A. Freeze dehydration

2.4 Food and self-connection

When individuals forage and hunt from a landscape, they find easily that they have an interest [in the health] of that landscape, and in the organisms that navigate that landscape. When someone has a relationship with a landscape, because one experiences the acquisition of food from that landscape, then one has a vested interest in that place. The same goes for the organisms themselves, the harvesting and hunting of a species gives the "predator" a relationship with that species that someone who thinks ill of harvesters and hunters might not have. There is a saying among traditional hunters:

> *"People who hunt deer love deer more than people who think hunters shouldn't hunt deer because it is not nice ... who think deer shouldn't be hunted. Hunters who hunt deer really come to know deer deeper, like they know themselves. Some people struggle psychologically with how ecology works, in that organisms eat other organisms. That is nature."*

Technology Support: Information Processing Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: information processing, information technology, computational servicing, information interfacing, information system production, information system servicing, data processing, computing services

Abstract

An information-based society relies on information, including data, flows, and processes, in order to reconfigure itself optimally. Information is "stored" in repositories, which take up physical space. Information in the form of signals can be transported between physical spaces. Information can be displayed to a user, it can be monitored for changes, or it can be kept within a closed loop framework. A society can model itself informationally, the most useful representation being a complete simulation. In a real-world information system, there is information that is purely conceptual, there is information about objects, and there is information about distances between objects. And, in an adaptive information system, there is past, present, and possible future, information. All material designs are conceptual information (ideas) prior to their creation in a shapeable [spatial] environment. The configuration of objects in an environment has consequence, the results of which may be accepted as data [feedback] into an organizing information system. Early 21st century information systems are sustained through combined hardware-software packages known as computers. Computers store data in matter, and process data through patterns of electrical powered matter configurations. Calculation upon all societal-level accountable information allows for the unified socio-technical, economic planning of society at a global level without trade or coercion. In a community-type society, working groups develop an information system standard, and the InterSystem/Habitat Information Team engineers the information systems service. Through decisioning processes, information that has been structured through some social ordering becomes integrated into solutions upon which teams take action. Actions have affects in a spatial environment; those effects are fed back as data. In a community-type society, all decisioning is transparent and must be understood prior to action. Information is a conceptual-type of access.

Graphical Abstract

Not Currently Applicable

A.k.a., Information cycling service.

No content here yet.

Technology Support: Communications Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: communications, communication system, communication service, language interface

Abstract

When signals are moved coherently, communication occurs and information becomes available as required. Communications refers to signal [electromagnetic object] generation and processing. For communications to occur, the entity receiving the signal must be able to interpret and/ or otherwise understand the signal. In order for entities in different locations to understand each other's signals, the subjects (entities) must have a common protocol for sharing and or otherwise generating meaningful signals (Read: object signal creation). To the entity that generates the signal and a receiver of the signal, the signal is a message (i.e., an information packet with meaningful data). Communicating systems use protocols for coherently sharing messages. At the software level, each subject-relevant meaning is encased in a technical level meaning that ensures data (meaning) is received without error (e.g., header and footer metadata). In order for sharing to occur, there must be a protocol used by all subjects and systems. That communications protocol provides pre-determined information flow responses based on provided information. A protocol defines the format of signals (messages), which is a language, though it is not commonly called that. Programming languages are those languages whose semantics describes computations. Protocols in general do not describe computations; though, some specific protocols do, such as, RPC protocols. A protocol uses well-specified signals to express well-specified meanings, but rather few of them (only two different meanings). Thinking languages (e.g., English, Mandarin, Russian, etc.) allow people to think about and to coordinate (or not) their work and service lives. Pilots and traffic controllers use protocols to communicate, some of which are highly complicated. Essentially, a protocol is a language and a language is a protocol, and subjects must share a protocol in order to think and communicate coherently. The communications service system is information oriented.

Graphical Abstract

Not Currently Applicable

No content here yet.

Technology Support: Transportation Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: transportation, transportation system, transportation service, distribution, distribution system, distribution service

Abstract

When objects are moved coherently, transportation occurs and access [to objects] becomes available as required. Transportation refers to object translocation (i.e., the movement of an object from one location to another). For transportation to occur, the entity receiving the object must be able to available to receive and to understand the object. In order for entities in different locations to understand each other's outputs, the subjects (entities) must have a common protocol for sharing and or otherwise generating meaning from objects. An object is a structure (shape) that may or may not have meaning to subjects. Transporting systems use hardware and software for optimizing the safe and effective transportation of objects. At the software level, each object is encased in a material enclosure to some varying degree, which transports the object without loss of the object. In order for the population-scale sharing of objects to occur, there must be a hardware transportation network signaled by environmental sensors upon which a software platform calculates the optimal packaging and logistics for transportation solutions. A transportation package moves through a network of well-specified signals-object processing devices that guide it to an intended destination. Within a habitat service system, these destinations can be integrated with the infrastructural environment. For instance, think of an integrated pneumatic distribution system in an office or at a bank teller, and then imagine such a system (or, its equivalent) at the city level. Transportation can be easily scheduled or requested on demand. Some transportation is of a higher priority than other transportation, for instance, a medical evaluation, or the introduction of medical personnel into an situation. The transportation service system is materially oriented. All transportation is integrated within and between cities. Consider the reduced transportation logistics when everything needed is in a life radius and is in walking distance.

Graphical Abstract

Not Currently Applicable

1 Introduction

No content here yet.

2 Packaging

It is possible to optimally pack and store objects given what is known and available. Packaging is the task of packing something for transportation or storage. Product packing may include, but is not necessarily limited to:

1. Product Information
   A. Assembly Diagrams
      1. Lists parts
      2. Lists required tools
      3. Describes assembly procedure
   B. Technical Specifications
      1. Lists relevant parts, materials, or compositions.
      2. Describes relevant specifications of the parts. What is relevant?
         i. Energy requirements
         ii. Care / Maintenance Requirements
         iii. Disposal / Upcycling guidelines
         iv. Operating Guidelines
2. Product Protection
   A. From environment
      1. Abrasion
      2. Corrosion
      3. Infection
      4. Deterioration
3. Product Storage
   A. Each piece has a place
      1. Place is determined by how it is accessed (How it is accessed means by whom and in what order)
4. Product Transportation
   A. Moving the product from point of manufacture to place of access / use
   B. Moving product from one point of access to another
   C. Moving product from a state of access to a state of suspended access (a.k.a. "storage") Why?
      1. Seasonality of suitable access environment
      2. Reaction to demand for access

Technology Support: Materialization Service System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

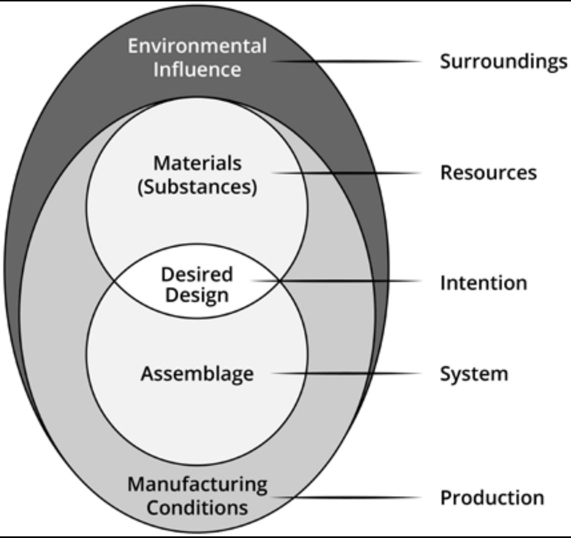
Keywords: materialization, material cycling material production, production, manufacturing, assembly, construction specification, material standardization

Abstract

There are objects that could exist, and there are objects do exist. There are intangible objects, such as the software model of a chair (i.e., the chair displayed on a monitor is an intangible object). When/if the chair is assembled to be a chair, then it will have been materialized [as an event, a new object-resource] in the system. A materialized chair is a tangible object. There is a distinction between objects recognizable on a monitor screen and made up of software, and a physicalized (materialized) chair that a subject can sit in. Any material environment must be composed of any one or combination of: life (subject), land, object, structure (architecture), and energetics (power). The prime objective of the materialization (materials cycling) system is to materialize and de-materialize objects, as compositions of resources. The decision system determines the solution for the next iteration of the societal configuration upon which teams may put effort. The human animal [subject] has evolved to send moral signals about their well-being and the state of its habitat, in three primary way: life, technology, and exploratory. Different types of societies could be looked at as different types of signaling configurations. A decision system allows for coherency in the fulfillment of societal configuration solutions to meet the demands of life, technology, and exploration opportunity. Different signal configurations may have significantly different effects on peoples actions, behaviors, and lives. The moral visions (values) people have for one another really matter from an evolutionary perspective. Humans are an evolved living individuated organism, and hence, have needs for life-support services at that habitat level. Humans explore life together through an exploratory service support system. Humans use technology to support their informational and spatial needs. This service system combines decisioning data with executionable data for InterSystem team tasking, which may combine in a variety of different ways depending upon a given society's configuration.

Graphical Abstract

Figure 20. *The reconfiguration of a material environment through intentional effort.*



1 Introduction

A.k.a., Materials cycling service system, the production and recycling system, the manufacturing system, fabrication, construction, production, generation, manufacturing, assembly, creation, generation, object realization and cycling.

There are many ways to bring new visualizations into real materialization. Humans have a need to cycle materials through their habitat(s). Some of the realized objects these materials take remain static and/or stationary for a long duration of time, relative to others. Materials cycle in and through habitats and human life systems. Materials become objects. Some objects and their relationships are of a sufficiently coherent nature to be called the materializations of life support, or technology support, or exploratory support. There are materializations that are appearing. There are materializations in memory of that which existed, and that which could exist (or, could be shown to potentially exist). When there is recognition that that which comes after can be different than that which came before allows for real-time materialization, then there is potential for adaptation. Engineering InterSysem teams work in the materialization of habitat service sub-systems which are constructed primarily as integrated city systems (a.k.a., total city system). Material flows can be traced to remain accountable for finite resources. To materialize is to be materially realized. In a community-type society, realization is executed (i.e., "is done") through team contribution. Teams contribute to the decided reconfiguration and operation of the habitat environment. Many objects become larger and more functional technological objects, which become useful services of function with object, for which there is a habitat of them.

To “produce” means to combine the materials (i.e., elements) and forces within an environment in coherent manner to change the expression (configuration) of existence. In simple words, it is possible to reconfigure the material environment to produce different objects and different functions. Here, “materialization” means to cause or become real, to become actualized in existence. Materialization also means to be realized or carried out. The materialization of the specifications for community, as a type of society, are partly experienced as a habitat service system and partly experienced as an information environment. Matter can be reconfigured. The intentional reconfiguring of matter can produce greater function (ability) within an environment. City systems allow for the localized coordination of materialized service systems. The idea of materialization carries with it the idea that embodied consciousness exists within an environment that is responsive to thought by some degree, for it is conscious thought that moves capable effectors in the system. Simply, a physical environment can respond more quickly or less quickly to though by embodied consciousness. A lower order dynamic of though responsiveness may be an environment where someone must think of vase (type of object) and then use their body to use a tool to produce the vase. A middle order dynamic may be an environment where someone must think of a vase and then use modeling software and 3d manufacturing hardware to produce a vase-like object. A higher order dynamic may be an environment that can communicated with via thought, such that the user simply thinks of a vase and the environment procedurally develops it and then 3D manufactures it in real-time.

Other words for the term “materialization” include, but are not limited to: production; manufacturing; construction; fabrication & prefabrication; and assembly. All of these terms include in their definitions the meaning that something occurs that changes or reconfigures the materially sensed expression or appearance of existence. For example, when you whittle wood you are materializing a different structure. When you construct a building you are doing similarly. When you assemble a toy you are modifying the position and location, and maybe even composition, of material for a new function. When you code a program you are assembling a new set of functions into the world.

Information takes material form and becomes active in terms of doing things. Nature doesn’t separate information and matter. When we understand this we can make objects that integrate with natural cycles. All production comes from an information system. Digital materialization is the term used to describe this, and it means a two way conversion between matter (material information) and information, or at least matter in different forms. One is more information friendly, electrons and photons, the other is static, matter. In order for this to exist we need: symbols; volumetric (3d or 4D space); constructive (modular); continuous (infinite surface, you can scale to any resolution you want); exact.

Fabrication is part of the production process in which an item is made (“fabricated”) from raw or semi-finished materials, instead of being “assembled” from already-made (i.e., pre-fabricated) components or parts.

Another less common definition which does not apply here is: “materialization” is the appearance of something out of nowhere, or from an unknown source.

INSIGHT: *For anything which is to be of service its formation must be accounted for.*

1.1 Manufacturing

NOTE: *Production and construction are essential to consider because they effects so many of the remaining choices (i.e., future probable choices).*

In generally, this document doe not use the term “manufacturing”, because that term is generally defined from a commercial perspective:

- “Manufacturing is the production of merchandise for use or sale using labour and ...”

1.2 Materialization phases and workflow models

A.k.a., Manufacturing phases, production workflow.

The phases of the production/materialization (i.e., manufacturing and integration) process are:

- Materials conversion > FAIT (Fabrication > Assembly > Integration > Testing)

Example with a simple window frame module:

1. Conversion = raw lumber; molten glass
2. Fabrication = frame wood; window panes
3. Assembly = assembled windows
4. Integration = window placed in building
5. Testing may be a separate phase, but may also be a process associated with each of the other phases.

1.3 Production order acronyms

Production order acronyms include, but are not limited to:

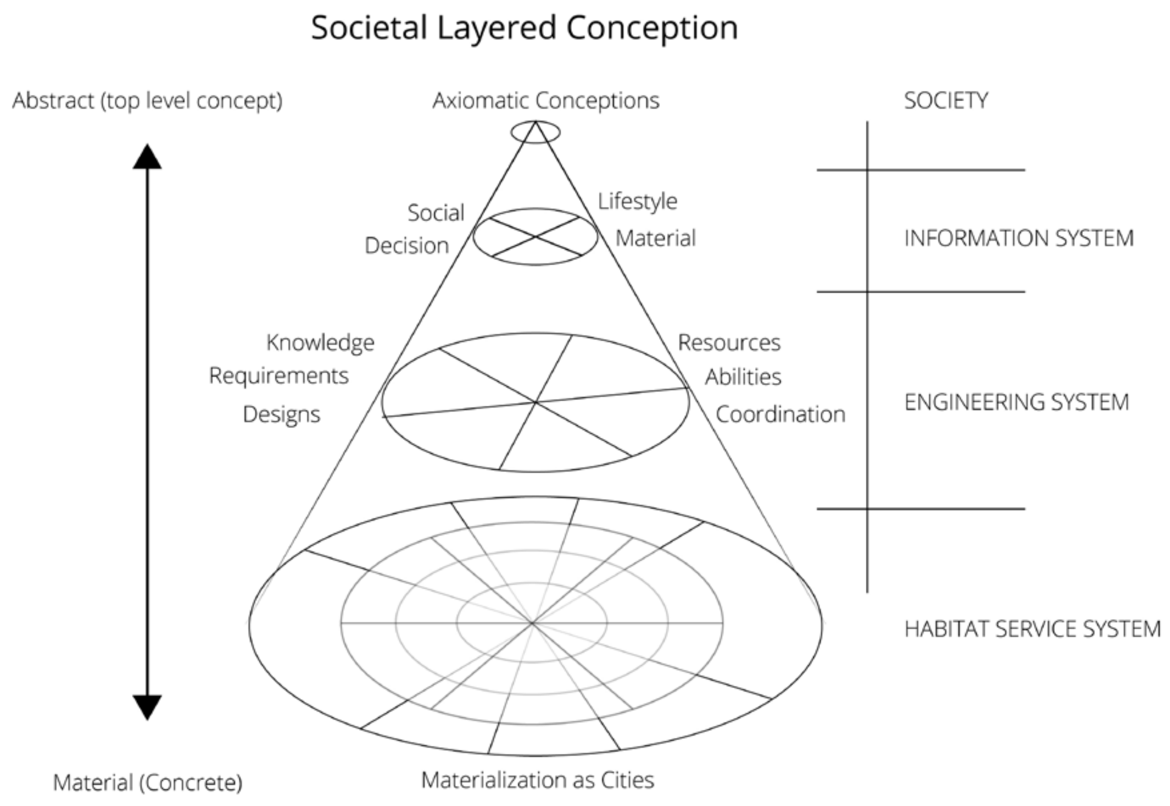
- ETO = engineer to order
- MTO = make to order
- ATO = assemble to order
- MTS = make to stock

1.4 Classic manufacturing processes

These are the classic manufacturing processes in commercial practice (in order of decreasing volume):

1. Continuous flow - large and continuously required production volume (e.g., oil and utilities).
2. Production line - high-volume production of standard products or “design window”.
3. Batch (high volume).
4. Batch (low volume).
5. Job shop - low volume, one-of-a-kind products.
6. Project - one of a kind or too large to be .

Figure 21. *The materialization of a society as a unified whole composed of a set of systems/dimensions representative of data (information processing), teamwork (the human effort), and physicality (the habitat operating system).*



1.5 Geometric modeling for 3D printing

The "geometric kernel" defines what is possible. It asks, how to capture and represent physical space and objects. It answers the question, "How do we define the world around us and how do we design?" A geometric model generates a modeling framework (describing the systems functionality), upon which exists a graphical and user interface. And, the graphical interface informs what is possible in terms of production and application simulation."

2 Materialization specification standards

- Master specifications (Masterspec®, SPECSystem™, MasterFormat™, SpecText®, BSD Speclink®, ezSPECS On-Line™, 20-20 CAP Studio, and many others)
- Local and national codes and ordinances
- Federal specifications (Specs-In-Tact, G.S.A., N.A.F.V.A.C., N.A.S.A.)
- National standards organizations such as the American National Standards Institute, National Institute of Building Sciences, the National Fire Protection Association, the National Institute of Standards and Technology, and the Association for Contract Textiles
- Manufacturers' industry associations (Fire Equipment Manufacturers' Association, American Plywood Association, The Brick Industry Association, etc.)
- Testing societies (American Society for Testing and Materials, American Society for Nondestructive Testing, Underwriters Laboratories)
- Manufacturers' catalogs (Sweet's Catalog File, Man-U-Spec, Spec-data)
- Industry-related magazines and publications (Construction Specifier, Architecture, Green Magazine On-line, Interior Design, Architectural Lighting, Architectural Record)

In the market-State, there exist a variety of materialization and other construction specification standards, including but not limited to:

- **American Institute of Architects (AIA)**, [aia.org] - publishes the AIA MasterSpec, which is produced for AIA by Deltek [deltek.com]. MasterSpec is a library of master specifications. MasterSpec is organized by MasterFormat number, but it fills in the actual specification content that MasterFormat isn't intended to provide.
  - *Student Access to MasterSpec*. (2018). Deltek. [avitru.com]
- **ASTM**, [astm.org] - ASTM International, formerly known as American Society for Testing and Materials, is an international standards organization that develops and publishes voluntary consensus technical standards for a wide range of materials, products, systems, and services. ASTM specifications and codes address material and testing procedures. The specifications issued by ASTM are organized on the basis of the type of material, and the letters prefixed to the specification number are indicative of the

material type. For example, letter A is for all ferrous materials; B is for all nonferrous materials, etc.
  - *Construction standards*. ASTM. Accessed: January 7, 2020. [astm.org]
  - *Additive Manufacturing Standards Activities*. ASTM. Accessed: January 7, 2020. [amcoe.org]
- **British Standards Institution (BSI)**, [bsigroup.com]
  - *Standards and schemes for certification*. BSI. Accessed: January 7, 2020. [bsigroup.com]
- **Building and Construction Authority of Singapore**, [bca.gov.sg]
  - *Publications*. Building and Construction Authority. Accessed: January 7, 2020. [bca.gov.sg]
  - *Guide on Construction of Industrial Developments in Singapore*. (2010). Building and Construction Authority. [bca.gov.sg]
- **Building Information Standards**
- **Construction Specification Institute (CSI)**, [csiresources.org] - publishes the CSI MasterFormat. MasterFormat is the intellectual property of the Construction Specifications Institute. MasterFormat is a numbering system for specifications. Construction specifications used in the United States typically conform to the guidelines of the Construction Specifications Institute's "MasterFormat". The MasterFormat index groups specification sections into identifiable disciplines using a six-digit system with digits in groups of two, such as: 01 24 30.
  - *MasterFormat Numbers and Titles.* (2016). [edmca.com]
  - *MasterFormat Specification Divisions (CURRENT). (2018).* ArchiMat. [archtoolbox.com]
  - *CSI 3-Part Formatted Specifications*. ARCAT. Accessed: January 7, 2020. [arcat.com]
  - *CSI MasterFormat™ Division List.* Builders Exchange of St. Paul. Accessed: Januayr 7, 2020. [plainsbuilders.com]
  - *Master Construction Specifications (PG-18-1).* (2019). US Department of Vetrans Affair, Office of construction and facilities management. [cfm.va.gov]
- **European Committee for Standardization (CEN)**, [cen.eu]
  - Search Standards. CEN. Accessed: January 7, 2020. [standards.cen.eu]
- **International Code Council (ICC)**, [iccsafe.org]
  - *Standard Development & Consensus Committees*. International Code Council. Accessed: January 7, 2020. [iccsafe.org]
- **International Construction Information Society (ICIS)**, [icis.org]
  - *ArchiveSpecifications*. ICIS. Accessed: January 7, 2020. [icis.org]
- **International Standards organization (ISO)**, [iso.org]
  - *ISO and Construction*. (2017). ISO. [iso.org]
- **U.S. Occupational Safety and Health Administration (OSHA)**, [osha.gov]
  - *Law and Regulations*. OSHA. Accessed: January 7, 2020. [osha.gov]

2.1 Construction specification standard types

The three most common types of building construction specifications used in commercial projects are:

1. **Performance specifications** - specification to address the operational requirements of an installation. The focus is on the project outcome, indicating how the final project must be able to function.
2. **Prescriptive specifications** - specification that contain detailed descriptions of what specific materials must be used as well as the installation instructions. This type of spec usually involves three key components:
   A. **General provisions:** requirements surrounding codes and standards.
   B. **Required products**: the type of products required based on performance and structural stipulations.
   C. **Execution procedures:** how to do the install and measure its effectiveness.
3. **Proprietary specifications -** specification that demand that only one specific product be used for a given installation. It is commonly utilized if the portion of a project requires a certain performance that only one product can achieve.

Specifications can also be classified according their allowance of one or more suppliers:

1. **Open specifications** - an open specification when an architect does not name a specific supplier or product and allows for substitutions to be made by the contractor. Open specifications do not limit competition, but rather is dictated by a set of standards that more than one manufacturer can meet, allowing for many alternatives to be submitted for approval. Performance specifications are often considered to be open.
2. **Closed specifications** - a closed specification lists specific products, systems and manufacturers, with no alternatives or mechanisms to apply a substitution. Closed specifications are most often seen when matching a specification to an

existing building, or when an exact duplication is important. By default, closed specifications are proprietary. However, they can be made "open" by not referring to a singular brand or providing requirements applicable only to a specific product. This is commonly done by adding "or equal" after the listed brand.

NOTE: *There are a variety of terms that relate to construction costing standards. For example, construction measurement standards refer to the way construction costs are calculated, classified, analysed and presented.*

2.2 Building design

In industry, building design is separated at a high-level into three categories (or disciplines):

1. Architecture
2. Structural engineering
3. MEP (mechanical, electrical, and plumbing)

These three disciplines all involve:

1. Connecting design to function.
2. Connecting design to fabrication.
3. Incorporating building data.

2.3 Surveying standards

- **International Federation of Surveyors (FIG)** [fig.net] - a federation of national member associations , FIG covers the whole range of professional fields within the global surveying, geomatics, geodesy and geo-information discipline.

3 The industrial-State construction process

The standard building process is: engage owner to determine requirements, pre-design (which includes programming, site analysis, etc.), then schematic design, design development, construction documents, and then contract administration during actual construction following documentation and procedures set by a larger authority. The abbreviations for these phases are generally: SD for schematic design; DD for development design; CD for construction documents, and CA for contract administration.

The design through to construction of an architectural structure in the market generally involves the following phases. From start to completion, the industrial-State construction process usually maintains the following phases:

Note: these phases may be repeated, skipped, or modified as necessary for each uniquely purchased construction project.

1. **Pre-design** - the project is defined by the customer in terms of its function, purpose, scope, size, and economics.
2. **Schematic design (SD)** services develop study drawings, documents, or other media that illustrate the concepts of the design and include spatial relationships, scale, and form for the owner to review. Schematic design also is the research phase of the project, when zoning requirements or jurisdictional restrictions are discovered and addressed. This phase produces a final schematic design, to which the owner agrees. The architect provides a preliminary evaluation of the program, schedule and construction budget developed in the pre-design phase and prepares a number of Schematic Design drawings illustrating the project to review with the owner. The designs lay out the program on the site and address schedule and construction budget requirements. The architect submits a preliminary estimate of construction cost to the owner.
    - Meet with client to determine design objectives, site conditions
    - Form design concepts and compare pros/cons of each
    - Designs incorporate preliminary code and materials research
    - Internal review
    - While client is reviewing design options, consultants such as interior designer, landscape architect, builder, manufacturers, engineers also review for added insight
3. **Design development (DD)** services use the initial

design documents from the schematic phase and take them one step further. This phase lays out mechanical, electrical, plumbing, structural, and architectural details. Typically referred to as DD, this phase results in drawings that often specify design elements such as material types and location of windows and doors. The level of detail provided in the DD phase is determined by the owner's request and the project requirements. The DD phase often ends with a formal presentation to, and approval by, the owner. Based upon the approved Schematic Design plans and required adjustments to program, budget and schedule, the architect prepares more detailed Design Development drawings describing the architectural, structural, mechanical and electrical systems, and makes adjustments to the preliminary estimate of construction cost.

- Refine client-selected schematic design
- Materials selection
- Determine alternates for design influeced by phasing or budget
- In-depth code research
- Incorporate consultant comments
- Develop construction details
- Internal review
- While client is reviewing final design, consultants review again
- Drawing may be sent for preliminary review by municipality or other appropriate authorities

4. **Construction documents (CDs)** - Once the owner and architect are satisfied with the documents produced during DD, the architect moves forward and produces drawings with greater detail. These drawings typically include specifications for construction details and materials. Once CDs are satisfactorily produced, the architect sends them to contractors for pricing or bidding, if part of the contract. The level of detail in CDs may vary depending on the owner's preference. If the CD set is not 100- percent complete, this is noted on the CD set when it is sent out for bid. This phase results in the contractors' final estimate of project costs. The construction document phase produces a set of drawings that include all pertinent information required for the contractor to price and build the project. Based upon the Design Development documents, the architect prepares bidding information, conditions of the contract, and an AIA agreement between owner and contractor. The architect advises the owner of adjustments to preliminary construction cost estimates and assists in filing documents for approval of governmental authorities.
   - Construction details are refined and specifications are finalized
   - Alternates are explicitly defined
   - Code analysis is generated, if applicable
   - Consultant drawings and documents are incorporated as appropriate
   - Internal review
   - While client is reviewing, consultants prepare final drawing tweaks for inclusion in construction set
   - Incorporate final client and consultant comments.
   - Final internal review
5. **Contract administration (CA)** services are rendered at the owner's discretion and are outlined in the owner-architect construction agreement. Different owner-architect contractor agreements require different levels of services on the architect's part. CA services begin with the initial contract for construction and terminate when the final certificate of payment is issued. The architect's core responsibility during this phase is to help the contractor to build the project as specified in the CDs as approved by the owner. Representing the owner, the architect observes the construction and administers the agreement between the owner and the contractor, determines that work is done in accordance with the contract documents, and certifies the contractor's pay applications. The architect reviews shop drawings, prepares change order documents, determines a date of substantial completion, and issues a final certificate for payment.
6. **Bidding & Negotiation** - The owner approves the Construction. Documents and the estimate of construction cost. The architect assists the owner in obtaining bids from General Contractors, negotiating proposals, and preparing and awarding contracts for construction.
7. **Construction phase** -
8. **Post-construction phase** -

Note: Architectural industry institutes notably state that pre-design and CA are not "basic services", but should be considered additional services.

In the construction industry there are a whole host of labor positions including architects, engineers, consultants, contractors, sub-contractors, and government agencies. The relationships therein can make a construction process a very tedious and challenging task to coordinate.

4 State jurisdictional building coding

A building code, or building control, is a set of rules that specify the minimum standards for constructed objects such as buildings and nonbuilding structures. The main purpose of building codes are to protect public health, safety and general welfare as they relate to the construction and occupancy of buildings and structures. The building code becomes law of a particular jurisdiction when formally enacted by the appropriate governmental or private authority. They exist to regulate business; they are regulation enacted through legislation.

What makes a good shelter? For thousands of years we've been attempting to regulate the idea. Even the bible had building codes. The book of Deuteronomy mandates railings on roofs to prevent falls. After the great London fire of 1666, cities began creating stricter regulations for how we build. Over the centuries the codebooks have only grown thicker. It's undeniable that building codes have saved lives, but something has changed over the past couple decades. Now thanks to the Internet the average homeowner has unlimited access to building techniques (both traditional and experimental) so we're experimenting more than ever, often without regard for the rulebooks.

Like most things in early 21st century society, codes have turned into big business. They are a key source of revenue for governments. They're also leveraged by many industries and organizations to push their political and economic agendas. These groups all benefit from code complexity.

These codes become the backbone of justification for the very existence of hundreds of Government building departments. With these departments come Plans Review, Zoning, Health and of course Permits and Inspection (code enforcement). This hierarchy, along with associated fees and expenses is, by design, supposed to protect human beings from harm, or to keep them safe from scientifically proven, dangerous building practices and or materials. Rules ,regulation, certification, licensing requirement, bonding and insurance requirements all work toward the common goal of "protecting and keeping safe" those that are not capable of doing so for themselves~ you really don't have a choice~ it's the law. Laws and ordinances are made as tools of enforcement to punish violators while keeping others in step. We should all be grateful for such a complex system of care and concern for the health and well being of human beings . After all, the majority of the human population is not capable of caring for or protecting itself from harm. People depend, even rely (for some reason) on this percieved service of care that is provided (at great cost) by the power of government. Apathy is nurtured as confidence as Oz grows. The people trust that they are being protected, but don't realize that this protection is based on the minimum requirement at every level. Large, "greed driven" mass housing developers, bankers, lawyers, realtors,building material manufacturers / suppliers and of course, the associated politicians recognized the obvious huge money opportunity that was tied directly to the " minimum code requirement" arrangement, and

naturally joined forces in producing the minimum standard in housing that was (and still is)allowable by laws, and selling, financing, insuring and appraising to the maximum based on geographic location.

In the market, without control, building can become a "free for all" where those who assume the labor role of "constructor" frequently display behavior indicating a lack of consideration for anyone but themselves and construct constructions that are dangerous or damage other constructions when doing their own. Others who have assumed the labor role of "constructor" may have the best intentions, but either forget something, or just don't know how to build something safely, and their construction fails and hurts something or someone.

Building codes reduce problems through coercion and force. In community, we design and test so that our active designs don't hurt us. Community is a place where everyone desires and decides to build safely. In community, we integrate our builds into a habitat service system designed to regeneratively fulfill our lives. In the market, builds are integrated into the market; and, they are be bought, sold and taxed.

The LEED Green Building Rating Systems are voluntary systems that assess the environmental performance of built projects across a spectrum of key criteria. From water and energy use efficiency to location, the impact of materials used, and more, LEED is intentionally designed to recognize buildings that go beyond minimum code compliance. While these minimums will vary from jurisdiction to jurisdiction, they should at the very least include the most current version of the model energy code as a mandatory minimum for all buildings.

In 2013, members of the US Green Building Council (USGBC) voted overwhelmingly (85 percent) to include Cradle to Cradle certification in LEED V4, which will even more stringently enforce the environmental qualities of materials used in green buildings, the opposite of what industry interests want. Now, those seeking LEED certification will get credits for Materials & Resources for disclosing and optimizing where building materials are sourced and purchased.

- LEED and Green Building Codes
- LEED Standard
- Green Building council

- **The International Green Construction Code (IgCC)**, [iccsafe.org] - The IgCC is a model code that provides minimum requirements to safeguard the environment, public health, safety and general welfare through the establishment of requirements that are intended to reduce the negative impacts and increase the positive impacts of the built

environment.
- IgCC: A member of the International Code Family. (2018). ASHRAE. [ashrae.org]

- **US Green Building Council (USGBC)**, [usgbc.org]
  - USGBC Publications. USGBC. Accessed: January 7, 2020. [usgbc.org]
- **World Green Building Council,** [worldgbc.org]
  - Green building rating tools. World Green Building Council. Accessed January 7, 2020. [worldgbc.org]
- **Cradle-to-Cradle (C2C),** [c2ccertified.org]
  - Cradle to Cradle Certified Version 4. (2019). Cradle-to-Cradle. [c2ccertified.org]

4.1 *Cradle-to-Cradle product standard*

> **NOTE:** *In the market all constructions are products.*

The Cradle-to-Cradle Certified Product Standard guides designers and manufacturers through a continual improvement process that looks at a product through five quality categories — material health, material re-utilization, renewable energy and carbon management, water stewardship, and social fairness. A product receives an achievement level in each category — Basic, Bronze, Silver, Gold, or Platinum — with the lowest achievement level representing the product's overall mark. Necessarily, Cradle to Cradle involves a labeling system and the declaration of ingredients.

Product assessments lead to the earning of a "certificate" indicating that the business standards body known as "Cradle to Cradle" has assessed and qualified the product to meet one of multiple achievements. Assessments are performed by a qualified independent organization trained by the Institute. The training and accompanying accreditation provided by Cradle to Cradle Products Innovation Institute comes at a financial cost (i.e., it is sold). The assessment provided by the "independent organizations" that purchased training from Cradle to Cradle comes at a financial cost to the product producer.

The idea that waste is actually food [as input for some other life process]. The essence of Cradle to Cradle is the importance of a closed loop, that only materials and processes that can be reused endlessly should be included in product design. And that if these guidelines are followed, humans can live in a world of abundance.

The cradle-to-cradle process accounts for the following:

- **Material health** - Knowing the chemical ingredients of every material in a product, and optimizing towards safer materials.
  - Identify materials as either biological or technical nutrients
  - Understand how chemical hazards combine with likely exposures to determine potential threats to human health and the environment certified
- **Material re-utilization** - Designing products made with materials that come from and can safely return to nature or industry.
  - Maximize the percentage of rapidly renewable materials or recycled content used in a product
  - Maximize the percentage of materials that can be safely reused, recycled, or composted at the product's end of use
  - Designate your product as technical (can safely return to industry) and/or biological (can safely return to nature)
- **Renewable energy & carbon management** - Envisioning a future in which all manufacturing is powered by 100% clean renewable energy.
  - Source renewable electricity and offset carbon emissions for the product's final manufacturing stage
- **Water stewardship** - Manage clean water as a precious resource and an essential human right.
  - Address local geographic and industry water impacts at each manufacturing facility
  - Identify, assess, and optimize any industrial chemicals in a facility's effluent
- **Social fairness** - Design operations to honor all people and natural systems affected by the creation, use, disposal or reuse of a product.
  - Use globally recognized resources to conduct self-assessments to identify local and supply chain issues and third party audits to assure optimal conditions
  - Make a positive difference in the lives of employees, and the local community

Cradle to Cradle certification launched in 2005 and rates products using five criteria:

1. Their use of environmentally safe and healthy materials
2. Materials are designed for recycling or composting at end of life
3. Manufacturing must make use of renewable energy and carbon management
4. Water stewardship
5. Social fairness

4.2 *The cradle-to-cradle red list*

In the cradle-to-cradle decisioning process, the "Red List" contains ingredients that inhibit all further forward movement. Or, it can move forward, but the company producing the Red Listed ingredient must be notified that the ingredient will be changed as soon as a suitable alternative becomes available. If a construction list has any of these ingredients, and production can't move forward without 1 of them, then the producer must send a letter to the provider of the Red Listed product stating that as soon as there is another option available "we wont use your product".

5 State jurisdictional permits

A.k.a., Construction permits, building permits, zoning permits, etc.

A "construction permit" (a.k.a., "building permit") is a permit/license required [as permission] in most jurisdictions for new construction. Building permits are to a large degree considered the profit end of building code, for they come with fees and taxes. The question is, Who can *issue (sell)* permits and who can *obtain (purchase)* permits? Note here that the word "obtain" is the word which jurisdictions often use when speaking of acquiring a permit. However, "obtain" is deceptive language -- what is really occurring is that the authority is forcing the purchase. Codes are enforced by the authority, and there is socio-economic punishment if they are not followed. The jurisdiction sells the permits, which have a financial cost associated with them. In most jurisdictions, the only person who can obtain a permit is a State licensed contractor, who pays the State for a license to do contract work in accordance with the standards (i.e., "building codes") of the State.

Book references

- Benyus, J.M. (2002). *Biomimicry: Innovation Inspired by Nature*. Harper Perennial.
- Baumeister, D., Smith, J., et al. (2014). *Biomimicry Resource Handbook: A Seed Bank of Best Practices*. CreateSpace Independent Publishing Platform.
- Linda, R.C. (2013). *The architects handbook of professional practice*. American Institute of Architects. Wiley.
- McDonough, W., Braungart, M. (2002). *Cradle to Cradle: Remaking the Way We Make Things*. North Point Press.
- McDonough, W., Braungart, M., Clinton, B. (2013). *The Upcycle: Beyond Sustainability--Designing for Abundance*. North Point Press.
- Rosen, H.J., Bennet, P.M. Harold J. (1979). *Construction materials and evaluation selection: a systematic approach*. Wiley.

Online references

- *Green building standards.* (2019). Harvard University. [green.harvard.edu]

TABLES

Table 37. **Technology Support > Materialization**: *Material cycling solutions.*

Material Cycling Solutions						
Repair & Maintenance	**Reuse & Distribution**	**Refurbishment & Re-manufacturing**	**Recycling**	**Cascading & Re-purposing**	**Organic Feedstock**	**Market**
Repair	Reuse, refurbish, maintain, redistribute	Re-manufacturing	Closed-loop production	Co-product generation from waste	Co-product generation from waste	Sell waste and/or product
Product life extension	Reuse	Upgrading	Re-materialization		Circular supplies	Waste exchange
Classic long-term model	Product life extension	Product life extension	Recycling and waste management		Resource recovery	Product lease

Exploratory Support: Scientific Discovery System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: scientific discovery, scientific experimentation, laboratories,

Abstract

Humans are, at least, adaptive and self-integrating systems who express evolved exploratory signaling ability. Young humans express exploratory behaviors oriented at learning (understanding) and controlling (testing). Every society is, in many ways, an experiment. Science involves environments appropriate for thinking and testing, for learning and demonstrating, for practicing and experimenting. It is favorable for humans to have the opportunity to select a scientific discovery system as one of its exploratory subsystems, because allows for safe navigation and technical validation. Some scientific discoveries may be operated at the community access scale and others may be operated through intersystem teams and working groups. Science enables reproducibility, and it is reproducibility that enables technology. Technology, then, enables societal-level services, and highly automated and coordinated fulfillment therefrom. New scientific discoveries may advance understanding and/or technological development. Safely coordinated discovery is possible at the societal level, when local individuals may choose to participate in scientific endeavours, particular as a life-learning experience. Some of these endeavours may be objectives at the societal level (e.g., the launching of a spaceship), and others may be localized, such as a youth running a common experiment in a lab in a city somewhere. The scientific discovery system is primarily run as a working group, but some cities and regions do have InterSystem habitat scientific discovery teams; those on other planets, for example. Some scientific discoveries are new and others are simply required for self-verification, since a community-type society is composed of self-directing individuals with life experience to be explored.

Graphical Abstract

Not Currently Applicable

Here, research may be more accurately known as scientific research. In the early 21st century, research is primarily an academician's category and development is an industrial category. Community organization, however, is not concerned with labor roles, but with where the actual information being worked on fits into a unified information model, and therein, a habitat service model. Scientific research and engineering development representing two distinct states of knowledge. Science is used to explore what happens and why in the natural world, while engineering uses the discoveries of science to create (and operate) technical constructions. Essentially, technical information about the operation of nature can be discovered [through science] and applied toward technical fulfillment [through engineering]. Science discovers what already is, and engineering creates that which isn't. Together, research and engineering/technology development (R&D) is composed of those activities which rely on science and engineering, and which are devoted to developing and operating socio-technical systems and services. Science informs, technology bridges. Technology is applied science. Scientists work with things that are testable, falsifiable, and reproducible. Then, systems engineers work through scientific discoveries to develop technologies.

The outputs of research feed into engineering development, and the output of engineering development are technical service modules (i.e., technologies) which are in use and operational as part of the habitat service system.

Science is about explaining and technology is about developing technical systems through trial and error.

Exploratory Support: Technology Development System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: technological development, technological engineering, laboratories

Abstract

Technology is the engineered interface of a commonly interrelated world. Technology enables function, as well as destruction of all function. Technological advancement continues over time to look like a more thought responsive environment. Humans can build and understand technical systems together. When humans are building technical systems together, they may rely on a pre-developed standard for building the new system. Working groups develop technological systems and their standards. Teams construct and operate the standards of a habitat service system. Technological development is an exploratory behavior, naturally enjoyable to humankind. Engineered technology is synonymous with high certainty deliverables, given the data and what is known. Technologies are operated. Some technologies are enjoyable by humans to operate. Technologies, or the use of technologies, that detract from human life or harm human operation should be reduced and monitored where necessary. An InterSystem Team herein is responsible for containing, sustaining, and restoring technological disasters (e.g., Chernobyl, Four Mile Island, etc.). This team would operate in conjunction with other InterSystem teams. In many ways, a technology is a control package influential at the information and/or spatial level. Technologies can be shared, communicated, transported, and materialized together, or not. Technologies enable the extension of function throughout an environment. Technologies and their applications can improve and degrade the quality of a life. Technological development is an exploratory working group and InterSystem team that develops technologies.

Graphical Abstract

Not Currently Applicable

No content here yet.

See also: The Scientific Research System.

Exploratory Support: Learning System

Travis A. Grant

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: Learning system, education service, learning lifestyle

Abstract

A learning system isn't so much an infrastructural integration, but an arrangement of services, some of which are informational and others of which are material that facilitate individual learning. Given that which is available, learning necessitates access to creative and collaborative technologies assistance (facilitation) in developing new skills, and mentorship in developing new established team member positions. Learning is a lifelong process that originates from within the individual. In a community-type society, learning about the society progresses best through participation in the society. Technologies and the flow state enable learning. More formalized learning is that which involves mentorship and possibly certification (e.g., like being certified to fly an airplane or conduct a surgical procedure). Learning requires materials that would otherwise be superfluous to a stable state society. The application of resources to learning structures facilitates societal adaptation and safety overall. The very nature of facilitation, is allowing another to determine what interests him (or her) to create or select their own "roadmap" for learning. There is structure, if desired. But, there is no such thing as a set curriculum, which must be completed within a certain period of time. Sharing one's own interests engages the interests of others. With proficiency come responsibility to encourage others who are less proficient and to use systems safely; not impress others, but express what is possible. Bringing out the best in others is how we find the best in ourselves. It is necessary to give to one another the tools to structure information in meaningful ways so that anyone can do something fulfilling with it. Failure is never a problem when there is play, the response may be. Knowing and experiencing are necessary. An event "happens" only when it is "experienced", and consciousness has been verified of its existence. A event is "experienced" only when it is "known", and consciousness only knows of its possible existence. All events are visualizable.

Graphical Abstract

Not Currently Applicable

No content here yet.

See also: The Lifestyle System Specification Standard.

Exploratory Support: Recreation System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: recreation, leisure,

Abstract

A recreation system is composed of many different components, the combination of which provide facilities and landscapes for indoor and outdoor recreation. Many entities are involved in the development and coordination of recreational areas and facilities within a habitat service system and its local network. There may be a recreational belt in any given city, and there may be recreational centers out in nature. In many market-State societies, a recreation system is a governmental function and a necessary State expense. In the market-State, recreation is largely considered to be unpaid/unpayable actions.

Graphical Abstract

Not Currently Applicable

No content here yet.

Exploratory Support: Art and Music System

Travis A. Grant

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: art, music

Abstract

An art and music system is composed of many different components, the combination of which provide facilities and tools for art and sound production. Many entities are involved in the development and coordination of art and music components within a habitat service system and local network.

Graphical Abstract

Not Currently Applicable

No content here yet.

Exploratory Support: Consciousness System

Travis A. Grant

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

Keywords: consciousness exploration system, spiritual service system

Abstract

A consciousness exploration system is composed of many different components, the combination of which provide facilities and tools for consciousness exploration, healing,a and spiritual connection/growth. Many entities are involved in the development and coordination of a consciousness service system for a habitat and local network.

Graphical Abstract

Not Currently Applicable

No content here yet.

Materials Accounting System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

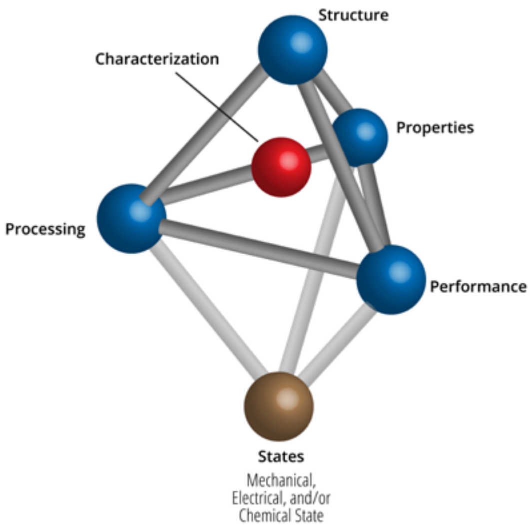
Keywords: material science, material science system, materials accounting, material flow, material composition

Abstract

A material system is a physical object in the world. A material system is naturally composed of materials (the synonyms of which are object, shape and surface). There are different material types, different material compositions, different material flows, and different methods of materializing. To intentionally reconfigure society at the material level, the materials of which society is composed must be engineered to meet societal requirements. When it is necessary to know the objects (of decisioning), then a materials system explication becomes essential to decision resolution. Materials can be combined and reconfiguring, and in doing so, their object properties can be changed (e.g., a neodymium iron core can be transformed into a magnet). A materials system is a repository of real objects (materials), their usages, and how they behave. There is an interrelationship among all objects in a material system. In some cases the tension of interrelationships is very weak, and at other times, there is a high tension. Materials selection is significant to solution determination. A complete accounting of material resources is essential if decisions are to be optimal.

Graphical Abstract

Figure 22. *Materials science characterization tetrahedron model shows how one aspect of a material affects the rest of the properties. If the structure is altered, the properties, characterization of atoms, processing, and performance will all be altered. Illustrates all the aspects of a material that an engineer should look for and understand.*



1 Introduction

The physical world is made of matter. All matter has physical presence, which may be sub-characterized as having 'object' and 'location' (i.e., distance from every object in the universe). When matter is used for human purposes, it is called a 'material'. Therein, materials science is the design and discovery of new materials.

1.1 Classification of matter

A.k.a., Classification of material objects.

All matter has the characteristics of:

1. State (a.k.a., phase)
   A. Solid
   B. Liquid
   C. Gas
   D. Plasma
   E. Bose-Einstein Condensate
2. Purity
   A. Pure substances
   B. Impure substances (a.k.a., mixtures)
3. Changes
   A. Physical
   B. Chemical
4. Properties
   A. Interaction properties
   B. Flow properties

1.1.1 Matter purity

In concern to purity, matter can be classified as:

1. **Pure substances** - matter that cannot be separated by physical means.
   A. **Element** - contains only one kind of atom. Elements cannot be chemically decomposed.
   B. **Compound** - contains two or more types of atoms in whole number ratios. Compounds can be chemically decomposed.
2. **Mixture (impure substance)** - matter that can be separated by physical means.
   A. **Homogeneous mixture (a.k.a., solution)** - uniform throughout.
   B. **Heterogeneous mixture** - non-uniform distribution.
      1. **Colloids** - Particle size: 1-1000 nm, dispersed; large molecules or aggregates; and particles do not settle out of the dispersing medium due to the effects of gravity.
      2. **Suspensions** - particle size is over 1000 nm, suspended; large particles or aggregates; and particles settle out of the dispersing medium due to the effects of gravity.

1.1.2 Material properties

A.k.a., Materials factoring, properties of matter.

Materials factoring includes:

1. Physical properties
   A. Chemical
   B. Electrical
   C. Magnetic
   D. Electromagnetic (a.k.a., optical)
2. Composition
3. Decomposition – including, wearing due to use and natural decomposition
4. Material resource flow
5. Contamination
6. Integration and De-integration

1.2 Matter sub-types

Additional matter types include, but are not limited to:

1. **Inorganic materials (inorganic compounds) -** most inorganic compounds do not contain carbon. Inorganic materials are generally derived from non-living sources, such as rocks or minerals, and encompass such categories as glass, ceramics, metals, minerals, clays, and metals.
2. **Organic materials (a.k.a., organic compounds, organic matrices)** - contain carbon and carbon-hydrogen bonds. They are solids composed of long molecular chains. For example, polymers, hydrogels, brushes, lipids, proteins, carbohydrates, nucleic acids.
3. **Hybrid materials** - incorporate both organic and inorganic constituents.
4. **Biological materials (a.k.a., biological compounds)** - self-organization of materials from the molecular level up. The basic building blocks are start with the amino acids and proceed to polypeptides, polysaccharides, and polypeptides–saccharides.
5. **Regenerative matter (a.k.a., living matter, biology)** - life and other biologically living materials.
6. **Elemental matter (a.k.a., non-living matter)** - the elements [of matter].
7. **Decaying matter**
   A. **Decomposing matter** - prior life; once living and now decaying/decomposing materials.
   B. **Radioactive decaying matter and transmutational matter** - the change of one element into another as a result of changes within the nucleus.
8. **Programmable matter** - matter with the ability

to change its physical properties (shape, density, moduli, conductivity, optical properties, etc.).

1.3 Characterization of materials

Generally, materials are characterized (compared) according to:

1. **Composition** - Composition tells what chemicals are in a sample. The most specific description will reveal the chemical elements that are present in the sample.
2. **Structure** - The three dimensional arrangement of atoms in a sample creates its structure.
   A. **Shape** (object).
   B. **Surface geometry** (surface physics and chemistry).
3. **Properties (physical and chemical)** - Properties are the observed characteristics of a sample.
   A. **Physical properties** - include how a material responds to mechanical forces, heat, and light, displacement, or a combination (e.g., density = mass / volume).
   A. **Chemical properties** - These describe what chemical reactions are likely to occur.
4. **Performance** - The performance of a material is discussed in the context of an application.
5. **Processing and synthesis** - Various methods can be used to create materials from existing substances. For example, processing a material could be as simple as hammering a piece of copper, or flaking arrowheads from a piece of flint. Synthesis implies a major change in chemical composition; for example, polymers are synthesized by cooking mixtures of chemicals, whereupon new molecular structures result.

These categories are useful ways to understand and sort different materials.

2 Material types

Traditionally, there are three basic material types (i.e., classes of materials) are:

1. Metals
2. Ceramics
3. Polymers

In material science, it is more appropriate to classify materials according to both their properties, but also their usages (*See below*).

2.1 Biomaterials

A biomaterial is any substance that has been engineered to interact with biological systems. Biomaterials are any synthetic or natural material used to improve or replace functionality in biological systems. Biomaterials are employed in components implanted into the human body for replacement of diseased or damaged body parts. These materials must not produce toxic substances and must be compatible with body tissues (i.e., must not cause adverse biological reactions). In other words, biomaterials are biocompatible and work synergistically with the biological host.

2.2 Bioelectronic materials

Bioelectronic use conductive polymers, organic semiconductors, carbon nanotubes, graphene, gold nanoparticles, photonic dyes, quantum dots, and microfluidic materials for applications in biosensing, bioimaging, wearable electronics, and implantable electronics.

2.3 Ceramics and oxide materials

Ceramics are inorganic non-metallic materials whose formation is due to the action of heat. A ceramic is a non-metallic material composed of inorganic molecules, normally prepared by heating a powder or slurry. Many common ceramics are made up of oxides or nitride compounds and are highly crystalline with long-range molecular order. Some ceramics are partially or fully amorphous, with no long-range molecular order; these are typically classified as glassy materials.

The six basic ceramic materials are (Note: the first five are classified as traditional ceramics and mainly made from natural raw materials):

1. Glasses (a.k.a., whitewares)
2. Clay products (e.g., brick and tile)
3. Refractories
4. Abrasives
5. Cements
6. Advanced ceramics - ceramics made from artificial

or chemically modified raw materials.

A. Electroceramics
   1. Electronic substrate, package ceramics
   2. Capacitor dielectric, piezoelectric ceramics
   3. Magnetic ceramics
   4. Optical ceramics
   5. Conductive ceramics
B. Advanced structural ceramics
   1. Nuclear ceramics
   2. Bioceramics
   3. Tribiological (wear-resistant) ceramics
   4. Vehicular ceramics

2.3.1 Glass

Glassy materials are hard, brittle, and noncrystalline. The lack of crystalline grains often results in optical transparency. The glass we are used to is a ceramic usually consisting of a mixture of silicates or sometimes borates or phosphates formed by fusion of silica or of oxides of boron or phosphorus with a flux and a stabilizer into a mass that cools to a rigid condition without crystallization.

2.4 Composite materials

Composites are mixtures of two or more bonded materials. Composites are the mixture of multiple materials, which in combination offer superior properties to the materials alone.

Composite materials include, but are not limited to:

1. Metal-ceramic composites
2. Metal-polymer composites
3. Ceramic-polymer composites
4. Concrete is a ceramic composite
5. Natural composites (e.g., wood, which is a composite material made from lignan and cellulose).

2.4.1 Concrete

Concrete is a ceramic composite made up of water, sand, gravel, crushed stone, and cement. The ingredients are mixed together thoroughly, and are poured into a form. After the concrete is completely dry, it has excellent compressive strength.

Concrete has excellent compressive strength, and unreinforced concrete blocks can be stacked miles high before the bottom-most blocks gets crushed. Concrete has little strength under tension. Modern builders work around this problem by making concrete into a composite, by embedding a rebar cage or mesh in a concrete slab, with enough thickness on either side so that when, under load, the armature stretches, the slab bends hardly at all. Because, if it did bend, cracks would instantly open up on the convex side, letting in moisture, causing the rebar to corrode, expand, and cause "spalling" (meaning the concrete structure falls apart). What's more, this is bound to happen eventually in any case, and so reinforced concrete slabs are engineered for eventual failure by being over-reinforced and under-cemented, because then they give warning of impending disaster in the form of cracks, as opposed to failing catastrophically.

Types of concrete include, but are not limited to:

1. Portland cement concrete (traditional 21st century concrete) - Portland cement is the source of the cement ("glue") that holds most modern concrete together. Making portand cement requires heating a mix of limestone and clays to 1,450C.
2. Ancient Roman concrete - Portland cement with a lyme and volcanic ash mixture. Portland cement is the modern type of cement. In seawater portland cement has a lifespan of ~50 years after which it corrodes. Ancient Roman cement lasts longer. The Romans perfected a mixture that used much less lime than portland cement and cemented at 900C or lower. The Romans mixed lime and volcanic rock for regular concrete structures, while underwater structures were made with lime and volcanic ash that formed a mortar. When this mix connected with seawater, a hot chemical reaction occurred that cemented the lime and ash mixture. The secret ingredient is aluminum-rich pozzolan ash and it turns out that oil-producing Saudi Arabia has a lot of it.
3. Aircrete - aircrete is made from a mix containing cement, lime and pulverised fuel ash (PFA) and a dash of aluminium powder. Aircrete is a material that combines the strength and durability of concrete which is physically light weight that helps make a home easy and fast to construct.
4. Cococrete/coco-peat - Cement and coconut fiber and lime and sand - a soft concrete that will start to moss up a bit when it gets wet, gives an ancient ruin look.
5. Bioconcrete - Concrete that heals itself using bacteria. The bioconcrete is mixed just like regular concrete, but with an extra ingredient -- the "healing agent." It remains intact during mixing, only dissolving and becoming active if the concrete cracks and water gets in. Tziviloglou et al., (2017) chose calcium lactate, setting the bacteria and calcium lactate into capsules made from biodegradable plastic and adding the capsules to the wet concrete mix.

2.4.2 Wood

Wood is a composite material made from lignin and cellulose. Wood makes use of a lignin matrix and cellulose fibers to form a polymer composite. The lignin holds the cellulose compressively in place so that the cellulose fibers can carry tensile loads. Wood has excellent structural properties, in light of its low weight and high strength.

2.5 Ceramic polymer

Ceramic with an inorganic polymer. So, it has the properties of a polymer in that it forms molecular bonds with metal and wood and themselves, as well as the properties of a ceramic in that they are highly crystalline (covalent and ionic bonding), and the properties of a cement such that it can be made into a powder and doesn't require high heat. Different aggregates can be added to the ceramics to make different qualities of building materials. If there is projectile damage, then ceramic can be mixed on site to patch damage. And, the ceramic can be sprayed to resurface the whole object.

2.5.1 Bioceramic polymer

Bioceramic are also known as chemically-bonded ceramics. Bioceramic is a chemically-bonded ceramic that forms strong molecular bonds like a polymer. Bioceramics can be chemically bonded with many different types of materials. Crucially, bioceramic has the same property that makes cement so useful: the ability to mix it into a slurry and pour it into a mold without using high heat. Bioceramic can be a highly environmentally friendly material. It can be easy to manufacture, and can be designed to be much stronger than concrete.

Bioceramic is fireproof, doesn't decompose readily when exposed to the elements (i.e., doesn't mold, doesn't rot, doesn't rust). Bioceramic can be made via carbon neutral processes. Bioceramic can be powderized before usage. Bioceramic can be used and built into multiple shapes.

Phosphate bioceramic polymers are some of the most usable and easy to work with bioceramic polymers.

2.6 Electronic / Optical

Electronic/optical materials are designed to conduct electricity or light. These materials may be metals, ceramics or polymers. These materials are carefully formulated to control the intensity, scattering, and bending of electrons or photons which pass through them.

2.7 Graphene materials

A.k.a., Carbon fiber.

Graphene is a polymer, an allotrope of pure carbon comprised of a single layer of atoms.

2.8 Metals, alloys, and magnetic materials

Metals are elemental substances that readily give up electrons to form metallic bonds and conduct electricity. Almost all metals have an orderly arrangement of atoms, resulting in a crystalline structure that may have multiple crystal phases bordering each other.

Some of the important basic properties of metals are:

1. Metals are usually good electrical and thermal conductors.
2. At ordinary temperature metals are usually solid.
3. To some extent metals are malleable and ductile.
4. The freshly cut surfaces of metals are lustrous.
5. When struck metal produce typical sound.
6. Most of the metals form alloys. When two or more pure metals are melted together to form a new metal whose properties are quite different from those of original metals, it is called an alloy.

Metals may be magnetic or non-magnetic. The magnetic properties of metallic materials are due to:

1. The atoms of which these metallic materials are composed.
2. The way in which these atoms are arranged in the space lattice.

Metallic materials are typically classified according to their use in engineering as under:

1. **Pure metals** - consist of a single element. Samples of these metals contain nothing but atoms of a single metallic substance.
2. **Alloys** - contain two or more elements or alloys melted and blended together, so their chemical formulas consist of more than one element.

2.8.1 Metal alloys

A.k.a., metallic alloys.

Metal alloys are classified as either ferrous or non-ferrous:

1. **Ferrous** - the group which contains mainly iron (Fe).
   A. Cast iron
   B. Steels
      1. Low alloy
      2. High alloy
2. **Non-ferrous** - other metallic materials containing no iron.

2.8.2 Metallurgy

A.k.a., Metallurgical science.

Metallurgy is the branch of science and technology concerned with the properties of metals and their production and purification. In other words, metallurgy, as a branch of engineering, is concerned with the production of metals and alloys, their adaptation to use, and their performance in service

Metallurgical science involves:

1. Physical metallurgy - the science of making useful products out of metals.
2. Process metallurgy (a.k.a., extraction metallurgy) - the practice of removing metals from an ore and refining the extracted raw metals into a purer form.

2.9 Metamaterials

A metamaterial (from the Greek word μετά meta, meaning "beyond" and the Latin word material, meaning "matter" or "material") is a material engineered to have a property that is not found in naturally occurring materials. Metamaterials are composite systems whose properties are dominated not by the individual atoms, but by the properties of larger, artificially produced structures or "meta-atoms." The concept of "meta" comes from the ability to engineer artificial materials, consisting of a composite of nanoscale structures, which can respond to other materials and to light in entirely new ways. A metamaterial is an engineered material specifically designed to exhibit a behavior that can only occur at specific organizations and sizes of materials. Metamaterials often seem to break the rules of physical behavior. In other words, metamaterials are composite media that can be engineered to exhibit unique electromagnetic properties. Simply, the field of metamaterials involves designing complicated, composite structures, some of which can manipulate electromagnetic waves in ways that are impossible in naturally occurring materials. Metamaterials are usually arranged in repeating patterns, at scales that are smaller than the wavelengths of the phenomena they influence.

Metamaterials derive their properties not from the properties of the base materials, but from their newly designed structures. Their precise shape, geometry, size, orientation, and arrangement gives them their new properties.

Metamaterials have even been shown to be capable of solving integral equations by encoding parameters into the properties of an incoming electromagnetic wave through a material structure that manipulates the wave in such a way that it exits encoded with the solution to a pre-set integral equation for that arbitrary input (Estakhri, 2019).

2.10 Nanomaterials

A.k.a., Nano-structured materials.

Nanomaterials, synthesized from both organic and inorganic materials, are defined as the functional [solid] materials with size below 100-nm in length along at least one dimension (Zhong, 2009). In other words, nano-structured materials are defined as solids having microstructural features in the range of 1–100 nm (nano = (1-100) × 10^{-9} m) in at least in one dimension. This includes both nano-objects, which are discrete pieces of material, and nanostructured materials, which have internal or surface structure on the nanoscale; a nanomaterial may be a member of both these categories.

2.11 Polymer materials

A.k.a., Plastics.

The term "polymer" is often used to describe plastics and other materials. Literally translated, polymer means "many units." These units are sometimes referred to as monomers, and they are the building blocks that form a plastic. Plastics/polymers are made up of millions of repeated links to make long molecules or networks that are tangled or crosslinked together. Almost all polymers use carbon atoms in very long chains. The carbon atoms may be attached to other carbon, oxygen, nitrogen, and hydrogen atoms. Polymers may or may not have an orderly arrangement of atoms. To form a plastic article, these monomers undergo a chemical change that causes them to become connected to each other. In addition to synthetic plastics, the term "polymer" also can be applied to natural biopolymers.

There are three basic types of polymers:

1. **Natural polymers** - polymers found in nature that have not undergone any chemical modification by humans.
2. **Biopolymers (a.k.a., natural biopolymers)** - natural polymers that have been chemically modified.
3. **Synthetic polymers (a.k.a., synthetic plastics)** - polymers that have been made entirely by humans.

Together, there are nine natural polymers and biopolymers:

1. Adhesion Proteins
2. Carbohydrates and Starches
3. Cellulose
4. Chitosan and Chitin
5. Dextrans
6. Gelatin
7. High-purity Collagen
8. Lignins
9. Polyamino Acids

Natural polymers tend to be readily biodegradable, although the rate of degradation is generally inversely proportional to the extent of chemical modification.

The polymer types by usage category include:

1. Biodegradable polymers
2. Block copolymers
3. Π-conjugated polymers
4. Dendrimers
5. Engineering polymers
6. Hydrophilic polymers
7. Hydrophobic polymers
8. Natural polymers and biopolymers
9. Poly(ethylene glycol) and poly(ethylene oxide)
10. Polymers for membranes
11. Polymer standards
12. Silicones

2.12 Semiconductors

Semiconductors are a special case of electronic material that combines two differently electrically conductive materials, usually ceramics. A semiconductor is also known as a P-N junction, where one material allows 'loose' electrons to move through an ordered structure, and the other allows holes (where an electron could be, but is not) to move in the same way. This behavior and the interactions between charge carriers and photons and phonons allows semiconductors to store binary information, form logic gates, and convert between voltage, light, heat, and force as sensors and emitters.

2.13 Textile materials

A textile is a flexible material consisting of a network of natural or artificial fibers (i.e., yarn or thread). Technically, a textile is an inhomogeneous porous medium. Yarn is produced by spinning raw fibres of materials to produce long strands. Textiles are formed by the following types of processes: weaving, knitting, crocheting, knotting, tatting, felting, braiding, etc.

2.13.1 Smart textiles

Smart textiles can be defined as textiles that are able to sense and respond to changes in their environment. They may be divided into two classes: passive and active smart textiles. Smart textiles can monitor an environment and be programmed to react in particular ways. (Koncar, 2016)

2.13.2 Classification of textile fibers

Classification of textile fibers includes, but is not limited to:

1. **Natural fiber**
   A. **Animal (protein derived)**
      1. Silk (from sericteries)
      2. Animal hair (from hair bulb)
         i. Alpaco
         ii. Cashmere
         iii. Camel
         iv. Feather
         v. Goat
         vi. Horse
         vii. Human
         i. Wool
   B. **Mineral**
      1. Asbestos
         i. Amosite
         ii. Crocidolite
         iii. Tremolite
         iv. Actinolite
         v. Anthophyllite
         vi. Chrysotile
   C. **Plant (cellulose/lignocellulose derived)**
      1. Seed fibers
         i. Cotton
         ii. Kapok
         iii. Loofah
         iv. Milk weed
      2. Bast fibers - Bast fibre (also called phloem fibre or skin fibre) is plant fibre collected from the phloem (the "inner bark", sometimes called "skin") or bast surrounding the stem of certain dicotyledonous plants.
         i. Jute
         ii. Flax (linen)
         iii. Hemp
         iv. Kenaf
         v. Kudzu
         vi. Mesta
         vii. Okra
         viii. Rattan
         ix. Ramie
         x. Rosella
         xi. Wisteria
      3. Leaf fibers
         i. Abaca
         ii. Agave
         iii. Banana
         iv. Fique
         v. Henequen
         vi. Manila
         vii. Raphia
         viii. Sansevieria cylindrica
         ix. Sansevieria ehrenbergii
         x. Sansevieria trifasciata
         xi. Sansevieria stuckyi
         xii. Sansevieria kirkii
         xiii. Sansevieria pinguicula
         xiv. Sisal
      4. Fruit
         i. Coir
         ii. Oil palm

5. Wood
   i. Soft wood
   ii. Hard wood
6. Stalk
   i. Rice
   ii. Wheat
   iii. Barley
   iv. Maize
   v. Oat
   vi. Rye
7. Grass/reeds
   i. Bamboo
   i. Bagasse
   i. Corn
   i. Sabai
   i. Rape
   i. Esparto
   i. Canary

2. **Human made (Manufactured)**
   A. **Natural polymer (artificial, regenerated)**
      1. Alzon (protein derived)
      2. Chitosan (natural sugars derived)
      3. Cupro
      4. Rayon (viscose/cuprammonium; cellulose derived)
      5. Modal
      6. Polynosic
      7. Deacetylated acetate (cellulose derived)
      8. Acetate (secondary triacetate; cellulose derived)
      9. Alginic (alginate)
      10. PLA (natural sugars derived)
      11. Lyocell (cellulose derived)
      12. Elastodiene
      13. Tencel
      14. Rubber (natural)
   B. **Synthetic organic (synthetic polymer)**
      1. Acrylic, polyvinyl
      2. Anidex
      3. Aramid/kevlar
      4. Carbon fiber
      5. Chlorofibre
      6. Elastin (elastoester)
      7. Fluoro fibre (fluoropolymer, teflon)
      8. Lastrile
      9. Melamine
      10. Modacrylic
      11. Novoloid
      12. Nitrile
      13. Polyamide (nylon)
      14. Polyester (aromatic polyester)
      15. Polyethylene
      16. Polypropylene
      17. Polyurethane
      18. Polyolefin (olefin)
      19. Rubber (synthetic)
      20. Saran
      21. Spandex
      22. Sulfur
      23. Trivinyl (vinyl)
      24. Vinyon
   C. **Inorganic**
      1. Metallic fiber
      2. Glass fiber
      3. Boron fiber
      4. Silica carbide

3 Material flows

A.k.a., Matter flows.

Material resource flows refers to the flow of matter/ materials within the physical (a.k.a., material) environment. Where resources are accounted for all identifying notations are tracked and calculated.

Important terminology in concern to materials flow includes:

- **Geology** - the study of how matter deforms and flows, including its elasticity, plasticity and viscosity. In geology, rheology is particularly important in studies of moving ice, water, salt and magma, as well as in studies of deforming rocks.
- **Geodynamics** - the deformation of earth materials.
- **Geomorphology** - that branch of earth science concerned with the shape of terrestrial surfaces.
- **Rheology** - the study of matter when it flows or is deformed.

3.1 Rheology

Rheology is the deformation and flow of matter; measured by a rheometer (or other). Rheology is otherwise defined as the study of flow behavior. Rheology is a well established area of study for a wide range of materials. In other words, rheology is concerned with the time-dependent deformation of bodies under the influence of applied stresses, both the magnitude and rate, whether the bodies be solid, liquid or gaseous. The term rheology originates from the Greek words 'rheo' translating as 'flow' and 'logia' meaning 'the study of', although as from the definition above, rheology is as much about the deformation of solid-like materials as it is about the flow of liquid-like materials and in particular deals with the behavior of complex viscoelastic materials that show properties of both solids and liquids in response to force, deformation and time. In practical application, rheology is most often applied to fluid materials (or materials that exhibit a time-dependent response to stress). In this sense, a secondary (or sub) definition of rheology is the study of the relationship between force (stress) and deformation (strain) of engineering materials under a set of loading and environmental conditions. (*A Basic Introduction to Rheology*, 2016)

Knowledge of rheological behavior is essential in numerous ceramic processing operations that involve slurries or pastes, including (Moreno, 2001):

1. Beneficiation (e.g., wet mixing and milling, atomization, and filtration).
2. Shape forming (e.g., slip casting-based methods, extrusion, roll forming, injection, and tape casting).
3. Coating/deposition (e.g., enameling, dipping, screening, printing, electrophoretic deposition, and spraying).

Rheometry is the method used to analyze the rheological behavior of a material. Rheological properties of a material are noted when a force is exerted on it, and as a result of which it deforms or flows. The extent to which a material deforms under a certain force depends strongly on its properties. Therein, rheometry refers to the experimental technique used to determine the rheological properties of materials.

Rheometers are measurement instruments for materials flow used to determine flow properties and viscoelastic properties of a material, the most notable of which is a:

- Rotational rheometer (a.k.a., viscometer) - measures shear flow and viscosity.

4 Materialization

Materialization is the process of acquiring, producing, and using materials.

4.1 Materialization calls

From a materials science perspective, the process of materialization involves the "call" of [factual] information from the following categories:

1. **Resource call:** Materials have properties.
2. **Material call:** Materials may be combined into material compositions to change the expression of [material] properties.
3. **Technology call:** Material compositions may be connected to perform technical functions as a technology module.
4. **System call (a.k.a., Service call):** Material configurations [as technology modules] may be integrated into service through a service system.

4.2 Materials quality control

There are significant differences between quality control in industry and in community. These differences are socio-economic in nature. In industry, the quality control process is as follows: people create things for another group of people; another group of people review and assess the first people's creations. In community, the quality control process is as follows: users openly and collaboratively create things for themselves while interfacing with a unified information model.

In industry, the outputs of processes must constantly be assessed and reviewed due to the presence of significant unknowns (i.e., due to the drive to conceal inputs, processes, and outputs for competitive advantage in the market). In community, the outputs of processes are significantly known due to the presence of a collaboratively developed, unified, and transparent information model. In community, inputs and processes and outputs are available for all to see, and for all to improve. In community, new outputs are tested prior to integration into the service system, whereupon they are tested again. And, once they are operational, we sense and otherwise monitor for signaled changes from the environment as feedback for improving and otherwise adapting (and evolving) our systems (i.e., our inputs, processes, and outputs).

In industry, independent reviews are essential. Competing entities are vying for finite market space, and so, "independent" entities are necessary to check the work of the other competing entities who are behaving for their own advantage. In community, all technical information is open source, and we recognize that we can behave in a way that is to everyone's advantage.

At the systems level, at the level of the Community's unified information model, we can all see and all check each others work. And, we work through our potential, purpose, and play. In industry, people work and compete for "income". In industry, due to the socio-economic consequences associated with reviews and assessments, they are generally taken as judgments. In community, the evolution of a model is seen as a benefit to all.

5 Biomimicry

Biomimicry is the study of the function of biological structures. Biomimicry takes design guidance from nature. We are now beginning to remember that other organisms are doing things very similar to what we need to do in ways that have allowed them to live gracefully on this planet for billions of years.

6 Material integrity

A.k.a., Material continuity.

Due to the entropic nature of material reality, all material that is formed into an intend functional system has a state of functional integrity (material change impacts function) and non-functional integrity (material change does not impact functionality).

6.1 Surfaces

A.k.a., Geometric surface.

Surfaces need to be carefully selected. Take a painted building for example. Would you rather see the buildup of "dirt" and get rid of it, or use a material that either did not allow for the buildup of dirt and/or did not show the buildup of dirt? Here, dirt is that which has been unintentionally added to or taken away from a surface (i.e., dirt is that which is out of place on a surface). So, dirt could be the buildup of particulate matter on a surface, or it could be deposits from the erosion/corrosion of the surface over time.

A surface has likely been selected because of its various properties. These properties are altered by dirt. Hence, the surface will need to be restored so that it is expressing the original properties desired of it, which may necessitate a [surface] **restoration cycle**.

> **NOTE:** *Losing function is losing the capacity to do something. As a space "wears", its structure can lose its capacity to carry on its function.*

6.2 Product expiration

Product expiration information tells the user that the producer cannot guarantee that the product will function as it is intended after the date (often, an estimation). Naturally, material compounds degrade with environmental exposure to conditions such as light, heat, and temperature. Some products have no expiration. Honey, for example, has would have no expiration date.

6.2.1 Expiration date

An expiration date or expiry date is a previously determined date after which something should no longer be used, or is not expected to function as expected. Expiration date is often abbreviated EXP or ED.

Scholarly references

- Balsubramanian, A. (2017). *Classification of materials*. Technical Report. DOI: 10.13140/RG.2.2.12792.34567 [researchgate.net]
- Estakhri, N.M., Edwards, B., Engheta, N. (2019). *Inverse-designed metastructures that solve equations*. Science, 363(6433), DOI: 10.1126/science.aaw2498

- Meyers, M.A., Chen, P-Y., Lin, A.Y-M., Seki, Y. (2008). *Biological materials: Structure and mechanical properties*. Progress in Materials Science, 53. [meyersgroup.ucsd.edu]
- Moreno, R. (2001). *Rheology*. Encyclopedia of Materials: Science and Technology. pp8192-8196. DOI: 10.1016/B0-08-043152-6/01468-6
- Tziviloglou, E., Wiktor, V., Jonkers, H.M., Schlangen, E. (2017). *Selection of Nutrient Used in Biogenic Healing Agent for Cementitious Materials*. Frontiers in Materials. 4:15. DOI:10.3389/fmats.2017.00015 [frontiersin.org]
- Zhong, W. (2009). *Nanomaterials in fluoresence-based biosensing*. Analytical and Bioanalytical Chemistry, 394, pp47–59.

Book references

- Koncar, V. (2016). *Smart textiles and their applications*. Elsevier.
- S. Karato. (2008). *Deformation of earth materials: an introduction to the rheology of solid earth*. Cambridge University Press.

Online references

- *A Basic Introduction to Rheology*. (2016). Malvern Instruments Worldwide. Whitepaper. [cdn.technologynetworks.com]
- *Types of materials*. Edmonds Community College. Accessed: January 17, 2020. [materialseducation.org]

TABLES

Table 38. Materials > Mechanics*: Major areas/branches of continuum mechanics.*

<table>
<tr><th colspan="4">Major Branches Of Continuum Mechanics</th></tr>
<tr><th>Type</th><th>Sub-Types</th><th colspan="2">Descriptions and studies</th></tr>
<tr><td rowspan="4">Continuum mechanics
The study of the physics of continuous materials</td><td rowspan="2">Solid mechanics
The study of the physics of continuous materials with a defined rest shape.</td><td colspan="2">Elasticity - Describes materials that return to their rest shape after applied stresses are removed.</td></tr>
<tr><td>Plasticity - Describes materials that permanently deform after a sufficient applied stress.</td><td rowspan="2">Rheology - The study of materials with both solid and fluid characteristics.</td></tr>
<tr><td rowspan="2">Fluid mechanics
The study of the physics of continuous materials that deform when subject to a force</td><td>Non-Newtonian fluids do not undergo strain rates proportional to the applied shear stress.</td></tr>
<tr><td colspan="2">Newtonian fluids undergo strain rates proportional to the applied shear stress.</td></tr>
</table>

Measurement Accounting System

Travis A. Grant,

Affiliation contacts: *trvsgrant@gmail.com*

Version Accepted: 8 June 2020

Acceptance Event: *Project coordinator acceptance*

Last Working Integration Point: *Project coordinator integration*

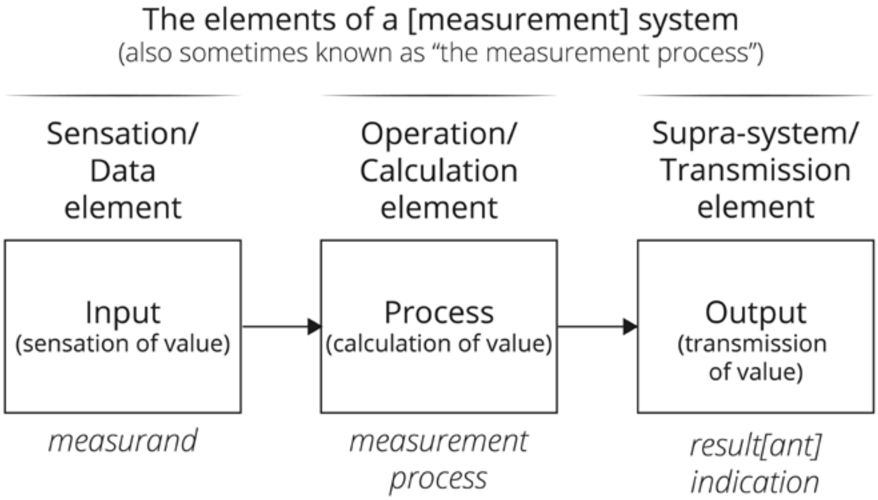
Keywords: measurement system, measurement service, metrological system, metrological service, measurement science, measurement engineering

Abstract

A material system is naturally composed of materials, which are the surfaces of objects. There are different material types, material compositions, material flows, and methods of materializing. To intentionally reconfigure society at the material level, the measurements between the components of which society is composed of at the physical level must be engineered to meet societal requirements. When it is necessary to know the distances between materials (objects, shapes), then measurement becomes significant.

Graphical Abstract

Figure 23. *The elements of a measurement system include sensation of input data, a means of determining alignment to prior sensed and/or prior synthesized values, and means of communicating the result of the alignment determination.*



1 Measurement in physical science and engineering

Science, engineering, and the material aspects of community are built upon measurement. Measurement is fundamental to scientific investigation and engineering. Hence, measurement is the foundation of science and knowledge. How well phenomena are measured affects what we know about them, and rigor in measurement increases the validity of analytical work.

Measurement is the foundation of scientific inquiry. In order to test hypotheses, theoretical concepts must be observed at the operational level. In simple terms, only that which is defined can be measured.

The physical task of designing and constructing an object (a.k.a., something) into the environment relies on measurement. In other words, engineers (i.e., individuals and systems that do these tasks) rely on measurement (and hence, metrology) to accurately design and develop physically functional systems. In order to have safe functioning of a material system it is essential for the systems design remain in some measured degree of alignment with the existent world and the principles of which it is composed.

In part, physics concerns observations, quantified through measurements, and expressed in units. The evolution of understanding around physical units is inevitably intertwined with a growing understanding of physics, the universe, and science itself. For example, after the introduction of the 'Celsius scale' (a concept based on the freezing and boiling points of water), it was only a matter of time before the notion of 'absolute zero' was conceptualized, and the 'Kelvin scale' was established. The Kelvin scale is based on the concept of "absolute zero". At "absolute zero", a hypothetical temperature, all molecular movement stops - all actual temperatures are above absolute zero. The kelvin scale has allowed for the measurement and construction of more complex[ly functional] technologies. Note here that the size of one kelvin degree is the same as the size of one degree Celsius.

Measurements are often associated with control or regulatory mechanisms. Therein, measurement allows for traceability and adaptive feedback. For example, in air-conditioning systems, temperature measurements determine whether heat flows are increased or decreased. In each case, the measurements precede decisions to increase or decrease, or to reject or accept.

All measurements may eventually contribute to a[n optimal] decision.

In order to communicate results unambiguously it is necessary for each of us to share the same scale for a quantity and to have access to the standards that define the scale. For metric scales the traceability problem is relatively simple: all measurements have to be related to a single standard. For the other scale types, the traceability problem can be more complicated because more standards are required.

Measurement is required if actualized (e.g., materialized) systems are to operate safely and remain in alignment with our highest fulfillment.

Ultimately, all measurements are used to help make decisions. Poor quality measurement data will result in inaccurate findings and faulty decisions.

> **APHORISM:** *To control the variable, it is first necessary to measure it. To measure the variable, it is first necessary to define it. To define the variable, it is first necessary to experience it. To experience the variable, it is first necessary to exist in a relationship.*

Measurement is required to validate hypothesis and attain understanding.

> *"When you can measure what you are speaking about, and can express it in numbers, you know something about it; but when you cannot measure it, cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind; it may be the beginnings of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be. So therefore, if science is measurement, then without metrology there can be no science." – Lord Kelvin*

1.1 Supra-system measurement objectives

In practice, a supra-system applies measurement for one purpose (that is, environmental information acquisition), and four main objects:

- **Discovery** – by a comparison of something of a presumably known quantity with something similar of an unknown quantity.
- **Diagnosis** – by a timely comparison of the actual quantity value with its normal range.
- **Alarm** – by continuously checking if the quantity value is inside or outside a parameter range.
- **Titration** – finely tuning an adaptive response action to bring a quantity value toward a targeted range.

Each objective type has different requirements, with the final intention off all objectives being to facilitate optimal decisioning, which together with an optimized information model, facilitate optimal living.

A self-organizing system uses the results of measurement to adapt itself to its environment and to improved functioning. The purpose of the measurement system is to link the observer to the phenomenological process (i.e., map/model for the observer the phenomenological process).

The intention of measurement is to acquire more information to inform a more informed and reliable representation of the real world. Wherein, measurement is used for (informs) the orderly and reliable representation of observation (i.e., measurement is the reliable, usable, and practical representation of reality).

1.2 Human decisioning and measurement

Humans desire material fulfillment, which comes [in part] from a specifically ordered approach to informing decisions:

1. **Quality [of material fulfillment]** - In order to generate and sustain fulfillment, it is necessary to take decisions.
2. **Valid decisions** - The optimal decisions cannot be made unless there are good numerical data on which to base those decisions.
3. **Correct numerical data** - Those numerical data, in turn, must come from measurements, which require accuracy [if they are to be useful in deciding optimally].
4. **Calibrated measurement (accurate measurements and calibrated instruments)** - The only way to get "good" numerical data is to make accurate measurements using calibrated instruments.
5. **Traceable standards** - If it is important to compare some [set of] measurements to other measurements made at other places and other times, the instruments must be calibrated using traceable standards.

1.3 The habitat service system measurement operational subsystem

Measurement is a system's process, because its results provide feedback that allow the supra-system to optimize its conditioned functioning. The material measurement system is part of the Community's core information system.

The measurement system integrates into the habitat service system as follows:

1. System: information system core
2. Sub-system: information acquisition and data processing
3. Operations: measurement (measuring), modeling, and calculating
4. Inputs: measurable information
5. Outputs: measurement models > procedures > scales > indications

The associated science and application of measurement is known as, metrology.

Every aspect (entity) of the habitat service system is measured for the optimization of our fulfillment and the well-being of our ecology. Each supra-system and sub-system involves measurement and all operational systems have performance/operational metrics. Therein, there are multiple sub-types of metrics including, but not necessarily limited to: community metrics, project metrics, product/service/process metrics, and quality metrics.

1.4 The function of measurement

Measurement allows for the design, development, and operation of technology. Engineered objects must be designed and developed within fairly narrow limits of tolerance it they are to work at all, particularly if operational systems require interchangeable parts. In this sense, measurement is a [conceptual] device for standardization, by which there is assurance of equivalences among objects.

A second function of measurement, one which shows its scientific importance, is to make possible more subtle discriminations and correspondingly more precise descriptions.

Unambiguously detailed classifications allow for a greater understanding of the principles of reality. Knowing that one thing "depends on" another is of incomparably less scientific worth than being able to say to just what extent changes in the first correspond to changes in the second.

Measurement applies to the process of inquiry in general the ability to verify, predict, and explain. In other words, measurements makes verification, prediction, and explanation possible (i.e., it makes science possible).

> **INSIGHT:** *Measurement is the comprehended awareness of a change from baseline (a reference).*

1.5 International measurement Standards

The International Vocabulary of Metrology (VIM, ISO/IEC Guide 99:2007) is the international standards document for metrology terminology, produced by the Bureau International des Poids et Measures. In general, a vocabulary is a "terminological dictionary which contains designations and definitions from one or more specific subject fields" (ISO 1087-1:2000, 3.7.2). The terminological vocabulary in the VIM pertains to metrology, the "science of measurement and its application". It also covers the basic principles governing quantities and units.

The VIM is now in its third edition (as VIM 3). VIM 1 and VIM 2 were mainly conceived by physicists and engineers for measurements in physics and engineering. Chemical measurement was considered to some degree in VIM 2.

The current ISO standards for measurement are:

- The International Vocabulary of Metrology (VIM, ISO/IEC Guide 99:2007)
  - *VIM3: International Vocabulary of Metrology*. (2017). Bureau International de Poids et Mesures. [bipm.org]
- ISO "Guide to the Expression of Uncertainty in

Measurement (GUM)
- *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. (2008). Bureau International de Poids et Mesures. [bipm.org]

1.6 The international standards definition of measurement

The current definition of measurement (in VIM 3) is:

> *Measurement is a process of experimentally obtaining one or more quantity value(s) that can reasonably be attributed to a [defined] quantity. Therein, measurement is the association of one or more numerical values to existent objects or events.*

Here, there are two principles:

1. Measurement is a process of attribution (to an earlier quantity or procedure).
2. The result of which is one or more quantity values and a measurement uncertainty).

Formerly (VIM 1), measurement was defined as:

> *A quantity subject to measurement. The measurement process is one of determining the value for the quantity.*

Per this definition of measurement there are two principles:

1. A determination process whose ideal outcome pre-exists the measurement itself.
2. The measurand has a single value.

1.7 Measurement sub-defined

Generally speaking, measurement is the process of observing, determining, and recording observations, in order to facilitate understanding and decisioning. That which is recorded by an observer becomes an input into a larger information system that a population uses to adapt. And yet, it is also general parlance for measurement to mean the activity of assigning a number to an observed relationship. Measurement is the assigning of numbers (or words) to empirical objects/events according in some definite organization[al model].

> **NOTE:** *Whereas measurement facilitates understanding [in part] by mathematical-statistical operations, decisioning is facilitated by the application of algorithms optimized for adaptive control.*

Measurement is a determination whose outcome pre-exists the measurement itself. The process determines (as in, "makes known") the value of some [existent] object or event (i.e., "thing") in relation to a pre-existing model of possible [finite or infinite] values, which are logically relatable.

Here, measurement is:

1. An inquiry into a fundamental domain of unitized information in the real [existent] world;
2. The result of which a greater understanding of the real world is programmed;
3. And hence, upon which a more optimized living environment is constructed.

Measurement is the usage of logical information acquisition and determination processes that experimentally interact with a quantity as the property of an entity that expresses some amount of an existence.

Therein, there is an order relation if some expression of a property can be greater or lesser (in quantity) than another. If there is an order relation, then there is a scale. Data with scale properties can be input into statistical operations to derive greater understanding and functionality from the surrounding world.

Measurement is the quantitative comparison between a predefined standard (or procedure) and a measurand to produce a measured result. Measurement is sometimes described (particularly in the social sciences) as the quantification of qualitative observation (of that which may or may not be observed).

Measurement involves 3 principal elements:

1. **The measurand** – quantity whose value is to be measured.
2. **The measurement system** – content, processes, and instruments for comparison.
3. **The observer or control unit** – the supra-system that uses and otherwise performs calculations on the data, and updates the measurement system information and available measurand classifications.

Measurement has two principal inquiry-problem categories:

1. Determining the value of an attribute/property of an entity.
2. Determining the class of entities to which the measurement relates.

1.8 The fundamental forms (types, procedures and operations) of measurement

All measurement takes one of three different fundamental forms (types, procedures and operations) of measurement. At the fundamental, level measurement is composed of the following axiomatic information processes: counting, ordering, and/or sorting. Each of the processes is itself a separate measurement process, although in a measurement system, these processes are generally combined. These processes may otherwise be called the [fundamental] procedures and/or operations of measurement.

The three fundamental forms (process types) of measurement are:

1. **Counting** – identifying [numerical] iterations of a similar [conceptual] pattern. This is by far the most commonly understood referent of the term 'measurement'.
2. **Ordering** – identifying [numerical] positions of a similar [conceptual] pattern. This is the process of arranging iterations into an ordered structure (i.e., arranging in order).
3. **Sorting** – arranging [conceptual and/or numerical] iterations of a similar [conceptual] pattern into categories representing sub-divisions.

These three forms are not just different versions of measurement. They are different fundamental types. The specific properties of each fundamental form of measurement determine:

1. The kinds of mathematical and statistical procedures that can be legitimately applied to a set of measures; and
2. The kinds of conclusions that can be meaningfully drawn from the application of the operations which have been run.

1.9 Characteristics of the conception of 'measurement'

'Measurement' can be viewed from the following perspectives and maintains (to a lesser or greater degree) the following characteristics:

1. **Measurement theory (a.k.a., the mathematical theory of measurement)** is the view that measurement is the mapping of qualitative empirical relations to relations among numbers (or other mathematical entities). The conditions under which relations among numbers (and other mathematical entities) can be used to express relations among objects. In measurement theory, and in mathematics generally, there is a scale of possible, increasingly complex, mathematical operations.
2. **Systems theory** views measurement as the acquisition of information from an environment (including, sub-systems) for adaptive and optimized functioning.
3. **Information theory** views measurement as the gathering, interpretation, interpolation, and integration of information about a system.
4. **Signal theory** views measurement is the reception of a signal from noise and the subsequent mapping of a new iteration to a pre-existing structure for understanding signaled iteration (i.e., sensation).

And, measurement can be viewed as having the following approach-oriented characteristics:

1. **Operational** – measurement is viewed as a set of operations (operational understandings) that shape the meaning and/or evolve the use of a quantity-term in the context of a larger intention.
2. **Empirical** – measurement is viewed as the estimation of mind-independent properties and/or relations.
3. **Analytical** – measurement is viewed as the discovered exploration of empirically operational patterns.
4. **Synthetical** – measurement is viewed as the comparative integration of a standard and a measurand, which produces a result upon which mathematical-statistical operations may be run to integrate new information about an existent [environmental] system.
5. **Model-based** – measurement is the coherent assignment of values to parameters in a theoretical and/or statistical model of a process. When measured parameters are numerical they are called "quantities". Here, measurement proceeds by representing the following interactions with a set of parameters, and assigning values to a subset of those parameters based on the results of the interactions:
   A. An object or event of interest.
   B. An instrument for measurement.
   C. An environment within which the measurement procedure occurs.

1.10 The 'determination' attribute of measurement

Measurement is, in part, a process of determination (or estimation. Measurement involves a determination (and/or estimation). Measurement is [in part] the estimation or determination of extent, dimension, or capacity (of a system), usually in relation to some standard and/or unit of measurement. The result of the process of measurement is the determination of a number of units of the standard (as a real number times a unit).

Among the attributes of measurement (i.e., primary characteristics or principles that compose the concept of measurement), 'evaluation' is one of those primary attributes. Here, evaluation refers to the processes of comparing, determining, and reporting a (counted) numerical value from a quantity-type source of information. The numerical value exists along a continuum of values. Here, previously unknown information is compared (viewed synchronously for pattern recognition) with a pre-existing [measurement] model [of referentially standardized, calibrated objects]. The value which is determined to map ("mirror") with the greatest alignment is then recorded into memory.

During these processes, the following events occur:

1. A numerical value (number) is determined.
2. That value is assigned to the quantity.
3. The number and quantity-type reference point are recorded into a memory.

Whereupon, a computational system (which may or may not be classified as part of the measurement system) initiates mathematical operations (as statistical calculations) on the data to acquire (determine) greater understanding (new information accurate to the accurate uncertainty of the data).

1.11 The 'mapping' attribute of measurement

Measurement is, in part, a mapping process; it is an activity of assigning a number or symbol to an entity in order to characterize a property of the entity according to given rules. In specific, measurement is an empirical to formal comparative mapping process that uses numbers. Measurement presumes that it is possible to preserve an empiric relation (connection with reality), using a numerical relation [known as a magnitude, quantity, or value). Measurement involves the application of 'number' to formally map empirical processes. Here, a number system facilitates real to abstract world transposition.

A number system allows for the empirical representation of real world patterns [of objects and relationships]. Measurement may be viewed as a mapping from the empirical world to the formal, abstract[ed] world. Here, the real world is the domain of the mapping, and the mathematical (or linguistic) world is the range.

For a measurement, there must be a corresponding numerical relation system, with symbols representing the entities and numerical relations corresponding to the empirical relations.

Mapping requires that a relationship shall have been established between the objects and the numbers so that each object there corresponds exactly one number, one point in the abstract space. However, in general, several objects may be mapped on onto the same point.

When the objects are so selected that the rule of assignment permits only one object to be mapped onto any point, then there is a on-to-one correspondence.

> **NOTE:** *The space into which objects are mapped need not consist of numbers. It would be more accurate to say that what is assigned to each object is a numeral rather than a number. The rule of assignment determines certain relationships among the numerals, and it is this pattern of relationships that constitutes the abstract space.*

This type of mapping follows a principle known as the 'representation condition'. The 'representation condition' states that a measurement mapping must map the entities into numbers and empirical relations into numerical relations in such a way that the empirical relations preserve, and are preserved by, the numerical relations. In other words, the relationships which exist between the attributes of objects in the "real world" are preserved in the numbers (or words) assigned these objects in the formal/abstract world.

There are two principal types of mapping processes:

1. Qualification is the mapping of observation to characterization.
2. Quantification is the mapping of observation to number.

1.11.1 Numbering in measurement

> **Insight:** *Numbers can be applied wherever there exists logic.*

Measurement consists of rules (applied logic) for assigning numbers (numbering) to attributes of objects. More specifically, measurement is the [logical] assigning of numbers to empirical events via the application of a set of rules (predefined rules that reference a standard rule). In essence, measurement is possible because of the syntax category (concept) of a 'number'. Numbers express [the presence of] delineation or iteration. In a sense, measurement is [in part] numerical input intended to map the delineation or iteration of observed relationships. In measurement, numerical inputs (numbers) represent a quantity (value or count) of entities in relationship.

In other words, measurement is the assignment of numbers to objects or events in a systematic manner. Or, said another way, measurement consists of rules for assigning numbers to attributes of objects/events. By definition, any set of rules for assigning numbers to attributes of objects is measurement. Measurement of some attribute of a pattern ("set of things") is the process of assigning numbers or other symbols to patterns ("things") in such a way that relationships of the numbers or symbols reflect the [real] relationships of the attribute being measured.

Measurement is the application of a mutually applied semantic system of numerical pattern recognition and categorization used to quantify a property (i.e., attribute, trait, or characteristic) of an existent system. A measurement, itself, takes the form of a number and accompanying unit that connects the number to a meaning, a significant event or concept. The number represents a comparison between the property of the system (or object) being measured, and the same property of a given 'unit of measure'.

Hence, measurement is the assigned estimation or determination of a number as a given 'unit' to a characteristic (property) of an existent object or event (represented as a concept), which can be compared through numbering to other objects or events (i.e., other concepts).

Measurement is the process of systematically assigning numbers to objects and their properties to facilitate the use of mathematics in studying and describing objects and their relationships. Measurement uses numbers to quantify - to process (transform) information into a [type of information known as a] 'quantity', so that mathematical logic may be applied. Numerical input allows for the logic and precision of mathematics (and hence, calculation) to be applied to the study of nature and the design of systems reproducing through its principles. Therein, measurement uses numbers to describe (real world) processes and events.

INSIGHT: *In order to coordinate resource flows in the material (physical) environment for human and ecological fulfillment, the material environment is initially understood to be composed of 'physical quantities', which logically, are quantized in some unit (by an axiomatic conception of existence).*

1.11.2 Mathematical integration and probability in measurement

Here, there is a pattern (variable), which has been separated into a sequence of sub-patterns (sub-set), and there is the probable recognition and integration of that pattern into an adaptive model of the original pattern, which can be described mathematically.

In mathematics, a 'measure' is a function that assigns a non-negative real number (or +(numeral infinity sign)) to (specific) subsets of a set variable (commonly represented as "X", "x", "x", "x", or "y", or possibly any other letter). This variable, the measure[-and], must be countably additive – the measure of a 'large' subset that can be decomposed into a finite (or countably infinite) number of 'smaller' divided subsets, is the sum of the measures of the "smaller" subsets.

INSIGHT: *Variables (e.g., x) are measurable functions, and units (e.g., mass, length) are measurable [real world] functions. There can also exist derived functions (e.g., power).*

In mathematics, the 'additivity' and 'sigma additivity' (a.k.a., 'countable additivity') of a function defined on subsets of a given set are abstractions of the intuitive properties of size (length, area, volume) of a set. Additivity is combinability (as in, the ability to [be] combined). The combing system is called 'add', and the process therein that does the combining (additivity) is called 'adding'.

In mathematical analysis, a 'measure' of a 'set' is a systematic process of assigning a number to each suitable subset of that set, intuitively interpreted as its 'size'. In relationship to visualization, a 'measure' is a combined supra-representation (conceptualization) of the [sub-]concepts of length, area, and volume. Here, parabolic>plane>solid ["Euclidean"] geometry is used to determine suitable subsets of the n-dimensional parabolic>plane>solid ["Euclidean"] space (Rn). Points in Rn are represented in coordinates as x = (x1, ..., xn), where x1, ... , xn are real numbers, and adding subscripts to a point in R n will always represent its coordinates.

There are four operational requirements that must be met for the combining of objects in measurement are:

1. **Commutative** – when two objects are combined the outcome must be the same regardless of which object is taken first.
2. **Associative** – the outcome must be the same regardless of how the combined objects are grouped – that is, the result of combining an object with the combination of two others must be the same as combining with the third the combination of the first two.
3. **Incremental** – the operation must be incremental with respect to the ordering of relation. If two objects are equivalent with respect to that relation, then the combination of either of them with some third object is no longer equivalent to the other one, but precedes it in the order established by the relation.
4. **Equalities** – if the two equivalent objects are each combined with objects equivalent to one another, the outcomes must be equivalent.

INSIGHT: *Our mapping of the underlying nature of reality is not discrete integers, but continuous functions.*

Here, measure theory is the formal model (and its underlying logical understanding) for how mathematics defines integration and probability:

1. **Integration (∫; in operation, ∫(x); a.k.a., function)** – measurable subsets are assigned numbers by [an operational] 'function'. The procedure of calculating an integral is called integration. An integral is a number associated with a function, and is usually called a "definite" integral. A "definite" integral is defined by a de-finiting (boundary or limiting) process. A definite integral is a formal calculation of area beneath a function. Integrals may represent the (signed) area of a region, the accumulated value

of a function changing over time, or the quantity of an item given its density.

A. The modern notation follows from Leibniz's notes, and given a real-valued function and real numbers, the definite integral is written:
 - ∫ab f(x) dx

B. Definite integrals have an indefinite form as well that serves as a partial inverse to differentiation. Just as differentiation measures a function's incremental changes, a definite integral attempts to "un-do" that. Hence, integrals focus on aggregation rather than change.

2. **Probability** – the measure assigned to the whole set is given the value, 1. Therein, measurable subsets are events whose probability is given by the measure. A probability measure is a measure with a total measure of one. A 'probability space' is a measure space with a probability measure. Every probability space gives rise to a measure which takes the value 1 on the whole space (and therefore takes all its values in the unit interval [0, 1]). Such a measure is called a probability measure.

A. A 'probability' or more precisely 'a finitely additive probability measure' is a nonnegative set function P(·) defined for sets A ∈ B that satisfies the following properties:
 - P(A) ≥ 0 for all A ∈ B, (1.2) P(Ω) = 1 and P(Φ) = 0.

1.12 The common parlance definition of measurement

NOTE: *In a mathematical operation, the input is an operand, and in a measurement operation the input is a measurand. The output of measurement is an operand value.*

In common parlance, measurement is the set of operations having the object of determining the "value" of a "quantity" of some "thing". Therein, the 'measurand' is that which is being measured; it is the quantity being measured. The result of a measurement [operation] is a value attributed to a 'measurand'.

Here, there are three important aspects of measurement not apparent from common parlance "definition" of measurement above that do apply to measurement:

1. In concern to number: The results of measurement need not be numeric: grade L, red, and carbon are all legitimate measurement results in the appropriate context. One of the most valuable aspects of symbolic representation is that the symbols in the models may be used to make predictions. Mathematical models and numeric symbols particularly help to quantify predictions that might otherwise be qualitative (or subjective).
2. In concern to intention - Every measurement has a purpose. This is the distinction between a meaningful measurement and meaningless assignment of numerals. In a great many measurements, especially outside the calibration laboratory, the purpose influences the design and outcome of the measurement. Consequently, measurement results may have meaning only within the context of that purpose. Results used for other purposes or gathered without purpose are potentially dangerous.
3. In concern to decisioning - Decisions are associated with real world consequences, which may be beneficial or not beneficial [to human fulfillment and ecological stability]. This highlights the need to know the uncertainty in a measurement in order to assess its applicable usefulness.

Measurement is the symbolic representation of existence to aid in understanding, adapting, and decisioning. It is the process of symbolically representing, organizing new information according to a pre-existing model (pattern) of information.

Any of the following could be symbolically represented through measurement (Read: including, but not limited to):

1. Concept, state, object, event.
2. Quantity, magnitude, amount, weight, degree, value.
3. Quality, property, attribute, characteristic.
4. Principle, rule, statement, argument, variable.

1.13 Conditions for measurement (measurability)

APHORISM: *Only quantity is measurable.*

Measurability can be understood in a number of different ways. Axiomatically, for any measurement, the characteristic (or property) to be measured is a quantity, in that it is an amount of something. Thus, it may be thought of as the sum of a number of elementary parts, or units, of that something. Here, measurement is equivalent to the counting of such units (with reference to a standard set of those units). From this analogy, it is possible to derive the conditions that must be met in order for measurement to make sense, that is, the conditions for measurability.

The minimal conditions of measurement are:

1. **A system of counting (applied pattern recognition)** - Counting is possible due to the

properties of natural numbers, which undergo an order, based on the relation "greater than or equal to," and may be added to each other. Counting is a way of assigning numbers to objects. The objects being measured are classes, and the individuals are numbered in order to be able to assign a measure to the class that they compose. Counting is a way of determining how many things there are of a certain kind.

2. **Empirical existence (ordered relation of experience as entity)** - Measurement implies the empirical existence of the entity for which some magnitude (count) may be specified. Measurement is not a thought experiment, it is empirical.
3. **Properties of existence (properties)** - What is measured is not an entity (e.g., a table or bird), but an property (the other type of entity) related to it (such as, its length or mass).

In concern to counting, objects can be counted, or ordered with respect to some attribute, does not suffice to enable the measurement of magnitude in such a way that arithmetical operations can be performed on the assigned numbers. Here, it is possible to answer questions of more or less, and even to determine how many objects in the field have a greater or smaller magnitude than some given object.

That which is measurable is:

1. Everything that is experienceable, or can be translated into experience.
2. Everything that is observable, or can be translated into observation.
3. Everything that is sensible, or can be translated into sensation.

Measurement may also be understood from perspective of an adaptive system – a system that uses the result of measurement to adapt its decided functioning. Therein, there exists the:

1. **The ability to measure (operational measurability)** - Operation pre-supposes functional design. Measurement as an operation (or series of operations) pre-supposes, at least:
   A. The method of comparison.
   B. The pattern for comparison.
   C. The procedure and apparatus used for obtaining the comparison must be provable.
   D. There are two operational requirements that must be met for measurement to occur:
      1. The standard (of reference) that is used for comparison must be accurately defined.
      2. There must be a pre-existing understanding (i.e, a model) to compare with that which is being measured.
2. **That which has the ability to be measured (empirical measurability)** - Measurability is an aspect of empirical properties (or, an empirical property), which allows for comparison with other empirical properties in terms of their ratio. The measurability of that which exists may be established by demonstrating ("proving") that:
   A. The characteristic under investigation involves an empirical order relation.
   B. Then, either:
      1. A physical addition operation allows the construction of a reference measurement scale and the performing of measurement by comparison with it.
      2. Or, by finding some physical law that allows the measure to be expressed as a function of other quantities.
3. **Utilization of the measured result (adaptational)** - For measurement to be of use (i.e., for the output to be useful for the larger system),
   A. The supra-system:
      1. There must be adaptive integration of control functionality. The larger system must be able to use the new information to change every aspect of itself and its decisioning.
   B. The measurement system:
      1. Validation must be acquired.
      2. Uncertainty must be accounted for.

Measurability may also be understood from the perspective of magnitude. Having magnitude (quantity) is sufficient for measurability. All quantities (including ordinal quantities) have magnitude. Note here that nominal properties do not have magnitude (and conversely to quantities that form scales, nominal properties cannot). Nominal properties do not have magnitude, and therefore, are not measurable; however, nominal properties are usable in a measurement system.

Finally, measurability can be viewed from two perspectives:

1. That which is conceptually "measurable" is a quality.
2. That which is numerically measurable is a quantity. And, a quantity is either a scalar or a vector.

1.14 Clarification of the term 'measure'

In common parlance, the word '"measurement" is used to refer to the result of a measurement process as indicated by a measuring instrument. In the science of measurement, this result is known as an 'indication', and not a "measurement" or a "measure". It is logical that the term 'measurement' relates to the whole [systems-oriented] process of obtaining a quantity value (indication) through comparison.

Measurement is sometimes defined as the act of determining a measure (quantity or quality) of some thing. Herein, a measure is a single quantitative attribute of an entity – the basic building block for a measurement.

Measurement (verb) is a form of observation-memory (verb), where the result of the observation-memory (verb) is the assignment of a quantity to a thing (the subject being observed). In common parlance, the term measurement may refer to the act of measuring (its verb form) or the result of that act (its noun form).

> **NOTE:** *The output of that which measure does is called an 'indication' (also sometimes called a measure, a measurement, or a signal response).*

To measure is to express as a number (or measure or quantity) an extent aspect or aspects of a physical and/or conceptual system.

A measure is a quantitative indication of the extent, amount, dimension, or size of some attribute of a system, product or process. A 'measure' is a quantity or amount given as a real number. It is the result of a method that involves an inquiry resolution process to determine how much [of something which is quantifiable] there is, or how many there are. A measure is a quantity logically assigned (given) to something (physical or conceptual) that can be quantified. Measurement (measuring and mensuration) is the act or process of assigning numbers to phenomena according to a rule "the measurements were carefully done"; "his mental measuring proved remarkably accurate".

There has long been confusion over the definition and appropriate usage of the term 'measure'. Although measurement is what something does, the term 'measure' has several meanings in common parlance. For this reason, it is generally not used without further qualification. For example, the term measure is often used in the following qualified ways:

1. An instrument of measure – an instrument, a device, a tool for determining measurements.
2. A measuring device - an instrument, a device, or a tool for determining measurements.
3. A unit of measure – a constant quantity that serves as a standard of measurement for some dimension.
4. A method of measure – the steps, stages, or processes taken to determine a measurement.
5. A scale of measure (level of measurement) - a classification that describes the nature of information within the numbers assigned to variables.
6. A particular measure[ment] – the 'indication', result, quantity value, or determined value of a measurement.

Expressions such as, "measuring a table", are not uncommon in parlance, though linguistically incorrect -- it is the [categorical] entity-property 'length' of the [categorical] entity-object 'table' that is being measured, not the 'table' as a category-entity itself.

To measure is to compare in a significant way any component (part) of a situation (system).

Here, complexity is expressed mathematically.

Note: In mathematics, a compound measure is a measure composed of two (or more) other measures (of a different type). A compound measure is based on two component measures (i.e., it is a measure with two integrated compounds; it is a "compound" measure). For example, speed is a compound measure composed of a measure of length (kilometers) and a measure of time (hours). Density is also a compound measure, composed of a measure of mass (grams) and a measure of volume (cubic centimeters). Density refers to how compact a substance is.

1.15 Clarification of the term 'metric'

In common parlance, the term metric has the following different, but related, meanings:

1. In general, metric means the whole conception and process of measurement. Etymologically, according to the Oxford dictionary, the word 'measure' is derived from the Latin word, mētīrī (or Ancient Greek, métron). Hence, the words measure and metric are often used synonymously/interchangeably. Here, the difference between metric and measure is:
   A. A 'measure' (dimension) is a fundamental or unit-specific term.
   B. A 'metric' can be derived from one or more measures (dimensions).
2. Tracking – In measurement, there is the tracking of that which is being measured over time. A 'metric' is a quantitative measure of the degree to which a system, component or process possesses a given attribute. A metric is a quantifiable measure that is used to track and assess the status of a specific process. Metrics are measures that are being tracked. There are two primary categories of tracked metrics:
   A. Performance metrics
   B. Calibration (Diagnostic) metrics
3. Standards – In measurement, there is the standardization of that which has been measured previously in time, for purposes of understanding, experimentation, and optimization. A 'metric' is a standard for comparison and/or reference. There are two primary categories of standard metrics:
   A. Performance standards
   B. Calibration (Diagnostic) standards

Wherever measurement occurs there may be a metric. In practice, metrics are the result of tracking measurements over time.

> **INSIGHT:** *'Normalization' occurs when metrics (moral and numerical) that [are known to] cause suffering, become normal, accepted and opted for by a population.*

In concern to tracking, a metric is a measure or combination of measures for quantitatively assessing, controlling, or improving a process, product, or team. Here, a metric is a standard reportable measure used to assess an operation.

In terms of performance, a metric is the desired and/or intended operating numerical value. A metric is a performance value ("performance measure") to be met by a system's process. New measurements (Read: measurement results) are compared to metrics (selected earlier measurement results, benchmarks). The metrics represent the decided and/or optimal value that the measure[d result] should be. A metric represents an earlier measurement(s) against which later (or newer) measurements will be taken to ensure that the system producing measured signals is operating as objective[ly] and functionally as intended. A metric is a previously defined value that a system, when measured, should express. It represents an earlier measurement used as a reference for later measurements. Here, 'metric' means the "standard" numerical or qualitative value, which should be output as a result of the measurement operation.

For any given project or system, where inputs, processes, and outputs are measured, each may have its own associated metric. For example, the inputs must be of a certain metric type (specification metrics); the processes have performance metrics (functional metrics); and, to the supra-system, the outputs have their own metrics usability metrics.

As a standard, a metric is a point of comparing or evaluating some property or attribute of existence and/or performance. A metric is a referential comparison standard against which some property, attribute, characteristic, or performance is being compared.

> **NOTE:** *In concern to measurement as a mapping process, the measurement mapping and rules are usually, together, called a metric.*

In terms of measurement, the standard to which the measurand is being compared is [called] a metric. For example, imagine the length of a solid object (measurand) being measured along ("against") a ruler. The ruler represents the measurement standard (i.e., the metric). And, that ruler was likely made from an earlier standard [metric]. If the ruler were a one meter standard ruler, then the metric [for measuring the object] would be a one meter standard ruler.

The process of defining new performance metrics involves, in order:

1. Determine entity category.
2. Identify measurement entity.
3. Identify attributes of the entity that are to be measured.
4. Define metrics.
5. This will define "success" or "failure" to meet a performance or other operations objective.

2 Metrology

CLARIFICATION: *Metrology should not be confused with meteorology, which is the science of weather phenomenon.*

In VIM3, metrology is defined as the science of measurement and its application. Metrology includes all theoretical and practical aspects of measurement, including the measurement of uncertainty and any field of application. It is the experimental and theoretical study of [weights and] measurement to ensure an optimized determination of the level of uncertainty in any field of science and technology. In practical terms, metrology ensures calibrated instruments deliver accurate results, and engineered systems operate effectively.

Metrology is an integral part of the theory of epistemology, gnoseology (Read: the study or philosophy of knowledge). Metrology is the study of obtaining accurate quantitative knowledge.

Metrology is the basis for empirical science and engineering. It allows for the generation of knowledge (as ordered information with logical uncertainty) of existence by transferring observational data into formal theory, and expressing them with logic (i.e., mathematical-statistical).

Performing a measurement means comparing an unknown physical, existent quantity (or quality) with a quantity (or quality) of the same type. The quantity of the same type to which the unknown quantity is being compared may be considered by a population as a reference, a standard, quantity. That standard may be expressed as itself, as in the case of a meter length ruler (a tool) for measuring a meter of length, or more complexly expressed as in the case of a magnetic resonance machine (an instrument) for measuring tissue position. A measurement necessarily involves a reference frame and therefore units. In the not so distant past, there were numerous units used to measure the same physical dimension, which caused engineering problems. The first coherent system of units only appeared with the French revolution: the metric system.

Metrology is a Greek language derived term for the science of measurement:

- Metro = measurement
- Logy = science (or, study of)

NOTE: *In its practical application, metrology requires standardization between cooperating individuals, groups, and systems.*

Processes in metrology include:

1. Establishing units of measurement.
2. Developing methods of measurement.
3. Analyzing accuracy.
4. Tracing measurements made in practice to reference standards.

There are two types of metrological study:

1. Scientific metrology refers to the inquiry, organization, and development of a measurement standard, and its revision.
2. Applied metrology refers to the adequate functioning of measurement instruments used in operational and testing processes.

At the base of metrology is the definition, realisation and dissemination of units of measurement. Properties of existence (i.e., an existent [real world] concept) are quantised by assigning a property value (i.e., a geometric numerical signifier, real number) in some multiple of a 'measurement unit'.

2.1 Metrological outputs

Metrological standards are the primary data output of metrological studies. The basic classification of measurement standards are:

1. The definition of a 'unit' type: based on some physical constant or an agreed-upon arbitrary standard. For example, in the case of a physical constant, the measurement of [the concept] temperature may be based on any of the following: absolute zero, the freezing point of water, the freezing point of oxygen, etc.
2. The realisation of the unit: by experimental methods and the scaling into multiples and submultiples, by establishment of primary standards. In some cases, an approximation is used, when the realisation of the units is less precise than other methods of generating a scale of the quantity in question. This is presently the situation for the electrical units in the SI, where voltage and resistance are defined in terms of the ampere, but are used in practice from realisations based on the Josephson effect and the quantised Hall effect.
3. The transfer of unit traceability: from the primary standards to secondary and working standards. This is achieved by calibration.

Standards are objects and/or relationships designated as used by all (i.e., "authoritative") for an acceptable and accepted reason [derived through axiomatic metrological concepts]. Whatever value they possess is useful for comparison to unknowns for the purpose of establishing or confirming an assigned value based on the standard. The design of this comparison process for measurements is metrology. The execution of measurement comparisons for the purpose of establishing the relationship between a standard and some other measuring device is calibration.

The ideal standard is independently reproducible without uncertainty. This is what the creators of the "metre" length standard were attempting to do in the 19th century when they defined a metre as one ten-millionth of the distance from the equator to one of the Earth's poles. It was later learned that the Earth's surface is an unreliable basis for a standard, as the Earth is not spherical and it is constantly changing in shape. But the special alloy metre bars that were created and accepted in that time period standardized international length measurement until the 1950s. Careful calibrations allowed tolerances as small as 10 parts per million to be distributed and reproduced in metrology laboratories worldwide, regardless of whether the rest of the metric system was implemented and in spite of the shortfalls of the metres original basis.

2.2 Metrology standard sub-types

There are three principal metrological standard sub-types in the production hierarchy:

1. **Primary standards**
   - Used for calibrating secondary standards.
   - At the highest level, a primary reference standard is assigned a value by direct comparison with the reference base.
   - International Prototype meter, Imperial Standard yard.
2. **Secondary standards**
   - Comparison for error correction between primary and secondary standards is continuous (or, as continuous as resources allow).
   - Exists as a secondary access control for reference of the primary references.
3. **Tertiary standards**
   - Exists as a tertiary access control for reference of the secondary references.

There are also:

1. **Working standards** – used by operators. Exist similar in design to primary, secondary, and tertiary standards. But, they are more numerous in access, and are made of easier to life-cycle (i.e., "lower grade") materials.
2. **Reference standards** – used for reference purposes.
3. **Calibration standards** – used for calibration of inspection and working standards.
4. **Inspection standards** – used by observing and analyzing (i.e., "inspecting") systems.

2.3 Modern standards for the dimensions of physical quantities

Currently, there are five independent units of measure (internationally recognized):

1. Temperature
2. Interval
3. Linear distance
4. Electrical current
5. Frequency
6. Mass

Any measurement can be based on one or more of these axiomatic units of measure (or, measurement units).

Pseudo-dimensional quantities involve angle (radian) measurement, of which there are two independent types:

1. Plane angle
2. Solid angle

**Note that a 'pseudo-dimension' is a dimension in which all tags are pairwise equivalent.*

Interested parties believe that eventually, standards organizations will define each of the independent units of measure in terms of the other four independent units. Length (metre) and time (second) are already connected this way.

It is probable that, eventually, all dimensional units of measure will be defined in terms of the other four [in] dependent units. Length, a linear distance measured commonly by the metre, and time, a frequency measured commonly by the second, are already connected this way. Linear distance can be measured using the known constant (or close to constant) speed of light, and hence, eliminate the metre bar artifact. And, time is measured by setting a cosmic linear distance as a reference standard.

NOTE: *Lesser known is the relationship between the luminance (candela) and current (ampere). The candela is defined in terms of the watt, which in turn derives from the ampere.*

2.4 The generation and application of metrological standards

In the market-State, the International Bureau of Weights and Measures (BIPM) develops measurement standards and enforces their application. In the United States, the National Institute of Standards and Technology (NIST) plays the dual role of maintaining and furthering metrology in the commercial and scientific fields. Presently, NIST does not enforce measurement accuracy directly. Instead, in the United States, the accuracy and traceability of commercial measurements is enforced per the laws of individual states. Therein, the government controls through regulation and enforcement of commercial measurement, as material sold by any unit of measure.

> **NOTE:** *Commercial metrology is also known as "weights and measures" and is essential to commerce of any kind above the pure barter level. Also note that the exact same term, "commercial metrology", is used to describe commercial calibration laboratories that are not owned by the companies they serve. In a commercial context, the term, "scientific metrology", addresses measurement phenomena not quantified in ordinary commerce. Calibration laboratories that serve scientific metrology are regulated as businesses only.*

In a hierarchy or market, a 'standard measure' (in this context, a measurement standard) is defined as something that is created, set up, and established as the norm by an authority as rule of the measure of quantity, weight, extent, value or quality. In other words, a body of people or systems in authority establish a set of rules for measuring things under their control and/or jurisdiction. However, 'standard measures' in community represent mutually integrated information that determines the resolution of a measurement-type inquiry or process. There is a difference between the market/state perception and the community perception, but the underlying concept that there shall exist a mutually used way to compare existence, remains the same.

In community, the metrology intersystem team resolves the determined "international" standards for measurement for the community, which is used by all community systems and sub-systems.

2.5 Axiomatic metrological conceptions

The axiomatic methodological conceptions in metrology (Read: concepts of or relating to the study of measurement methods) are:

1. **Accuracy** – Degree of exactness with which the final product corresponds to the measurement standard. How close is the observed measure to the actual (or, accepted) value. The measuring instrument/tool is a variable. Accuracy is calculated by the formula: % Error = (measured value – actual value) x 100 / actual value
2. **Requirements for accuracy** – what is needed in order to acquire a set degree of accuracy.
3. **Precision** – Ability to produce a measurement consistently. How finely tuned a measurement is, or how close multiple measurements can be to each other? The measuring instrument/tool is a variable. Precision is determined by the number of relative significant digits.
4. **Reliability** - Consistency of accurate results over consecutive measurements over time.
5. **Calibration** - The transfer of traceability from the primary standards to secondary and working standards is accomplished by calibration.
6. **Response time** – the time a system or functional unit takes to react to a given input.
7. **Traceability** - Ongoing validations that the measurement of the final product conforms to the original standard of measurement, and all calibrations therein are precise. Ongoing validations that the measurement of the final product conforms to the original standard of measurement

> **NOTE:** *Accuracy and precision may be demonstrated by shooting at a target. Accuracy is represented by hitting the center circle (the accepted/actual value). Precision is represented by the tight grouping of shots (they are finely tuned).*

2.6 Methods of measurement

The following are the most common methods of measurement in metrology:

1. **Precision or direct method** - measurements are directly obtained through . For example, micrometers, Vernier instruments, scales, and dial gauges.
2. **Indirect method** – calculation is used to visualize the measurement. For example, weight is length x width x height x density.
3. **Comparative method** – two measured values are compared.
4. **Coincidence method** – measurements coincide with certain lines and signals.
5. **Fundamental method** – measuring a quantity directly in related with the definition of that quantity.
6. **Contact method** – sensor/measuring tip touch the surface area.
7. **Complementary method** – the value of a quantity to be measured is combined with a known value of the same quantity. For example, volume determination by liquid displacement.
8. **Deflection method** – the value to be measured is directly indicated by a deflection of a pointer. For instance, pressure measurement.

2.7 Applied size categories

1. **Nominal size** – is the size of a part specified in the drawing. Note that nominal and basic size are often the same.
2. **Basic size** – is the size of a part to which all limits of variation are determined. Or it is the theoretical

size from which limits of size are derived by the application of allowances and tolerances.

3. **Actual size** – is the actual measured dimension of a part.
4. **Tolerance** – the total amount that a specified dimension is permitted to vary. It is the difference between the maximum and minimum limits for the dimension. A tolerance is the total permissible variation from the specified basic size of the part.
   A. Upper deviation (maximum, max)
   B. Lower deviation (minimum, min)

2.8 Metrological standards of measurement

There are perceptions through which that which is being observed and analyzed (i.e., measured) may be understood. These perceptions represented a scale of how fully the operation of the universe is understood.

1. **Line standard** – a distance, a "meter", is defined as the distance between scribed lines on a bar of metal under certain conditions of 'temperature' and 'support'. The meter, for instance, is the distance between the center portions of two lines engraved on the polished surface of a bar of pure platinum-iridium alloy (90% platinum and 10% iridium).
   - **End standard** – is expressed as the distance between two surfaces; generally, with the usage of a precision measuring mechanism (a measuring instrument). Dimensional tolerance as small as 0.005mm can be obtained. These, are not subject to ocular parallax effect because the instrument resolves the distance.
2. **Wavelength standard** – a "meter" is defined as the study and design of interferometry:
   A. Interferometry is a family of techniques in which waves, usually electromagnetic, are superimposed in order to extract information. It is the study and design of system that can account for and control the vibration of a medium as a rate of induction.
   B. The emitted/inducted wavelength of the cadmium line ($\lambda \approx 644$ nm), led to the definition of the ångström as a secondary unit of length for spectroscopic measurements. The ångström or angstrom is a unit of length equal to 10^{-10} m (one ten-billionth of a metre) or 0.1 nanometre.
   C. Krypton-86 ($\lambda \approx 606$ nm) was selected (in 1960) as the new wavelength standard for the [1] meter distance. Hence, the metre is the length equal to 1 650 763.73 wavelengths in vacuum of the radiation corresponding to the transition between the levels 2p10 and 5d5 of the krypton 86 atom.

2.9 Computational metrology and geometry

Fitting is the computational metrological term for associating ideal geometric forms to a discrete set of points sampled on a computationally manufactured surface.

1. **Datum establishment** – for relative positioning of geometric objects.
2. **Deviation assessment** – how far has a part deviated from its intended idea form?
3. Form tolerances (syntax and semantics)

Fitting is an optimization problem:

- Given a set of points X, fit ideal geometric element(s) Y that minimize an objective function involving distances between X and Y, subject to certain constraints.

The two principal types of fitting are:

1. Continuous optimization (e.g., least squares fitting)
2. Combinatorial optimization (e.g., minimax fitting)

Two popular fits:

1. Least Squares Fit – when the objective function uses L2 norm.
2. Chebyshev Fit – when the objective function uses L_∞ or other norm.

Conversely, the main purpose of filtering is to extract scale dependent information, and no compression of data. Filtering refers to convolution:

1. Convolutions of functions (e.g., Gaussian filters)
2. Convolutions of sets (e.g., envelope filters using Minkowski sums)

3 The measurement [comparison] process

DEFINITION: *The process that measures a quantity is known as a 'measurement procedure'.*

In common parlance, the term 'measurement process' could be used to refer to: (1) the measurement system as a whole, including its inputs and outputs; (2) everything included in 1, and the total conversion process of converting the source of the measurand into something measurable; or, 3) it could be used to refer solely to the measurement systems operational process(es). In other words, the processes which might be present to convert some object into an intended measurable constituent are generally considered to be part of the measurement process itself. In some cases, there may be a particular sampling procedure included in the process. In all cases, measurement results are obtained by performing measurement actions.

As the operational element of a system, the measurement process involves a set of operations having the object[ive] of experimentally determining the value of a [unique input] quantity, for a given attribute/ property of an entity, through observation (and hence, attribution) of its relationship to an earlier quantity. Therein, measurement is the process of assigning to some specific instance (of a quality, categorical property or attribute of existence), a numerical value (quantification) and a referential standard (unit).

NOTE: *A 'characteristic' (unique inherent quality) of a system is called a 'property'.*

Measurement (i.e., the measurement process) involves a series of actions (steps, stages) that take place in a defined manner. Some measurements are a single step, and others have many stages.

The purpose of the measurement process is to acquire new information (as comparative data) on empirical phenomena.

3.1 Conceptual phases of the measurement process

The generalized measurement process may be perceived to have the following conceptual phases:

1. Select an observable/sensible [empirical] event (or object).
   - Define the measurand by defining that which has an existent quality or quantity for which information can be acquired (or collected).
2. Develop a set of mapping rules (i.e., a scheme of principles for assigning numbers).
   - Define a standard comparison model.
3. Apply the mapping rule to each observation of that event.
   - Assign a number to a quality (property or characteristic) of an object or event, which can be compared with other objects or events.

3.1.1 Comparison inputs

INSIGHT: *Measurement is an information acquisition by a process of comparison.*

Measurement is the act[ion] or the result of a quantitative comparison between a predefined standard (procedure and/or model), and an unknown magnitude.

In order to complete the mapping, process a determined comparison must occur between [at least] two comparatively aligning inputs:

1. The measurand (unknown magnitude) - Some "thing" defined to exist from which more information can be acquired (the 'measurand'). The 'measurand' is the thing that is being measured, and for which a value will be determined. The measurand has a single value.
2. The comparative standard - A[n agreed upon] 'standard' method[ological scheme] of reference to determine the [standardized] value of the 'measurand'. The measurement standard (or standard of reference) is the pre-existing referential process and/or configured objects used in the comparison.

The value of the 'measurand' (i.e., the quantity value assigned to the measurand) is determined by its relationship (position and/or alignment) to the 'standard'. In usage, the standard [method or tool] of reference is used by an observer (or other decision processing, comparison resolution system) to assign a [quantity] value to the measurand by comparison with the standard in some pre-defined logical way (i.e., method or process).

3.1.2 Comparison methods

TERMINOLOGY: *A* **reference quantity** *value is a quantity value used as a basis for comparison with values of quantities of the same kind.*

There are two types of methodological comparison, direct comparison and indirect comparison:

1. **Direct comparison method** - Direct comparison with either a primary or a secondary standard. The direct comparison method involves a comparison of a measurand with either a primary or a secondary standard, which has the same physical nature as the measurand.
2. **Indirect comparison method** - Indirect comparison with a standard through the use of a calibrated system. Here, an empirical relation is

established between the measurement actually made and the results that are desired. The indirect comparison method is the main method that is widely used in contemporary measurement and control systems. The indirect method of measurement consists of two stages. The first stage involves converting both the standard and measurand into the type of output parameters that are convenient for further processing. The most common output parameters are electrical signals. The second stage of measurement provides a comparison of the first stage output parameters related to the standard and measurand.

3.1.3 Standard [of reference]

A [standard of] reference can be a measurement unit, measurement procedure, a reference material, or a combination.

1. **Reference material** – a sufficiently homogeneous and stable material with reference to specified properties, which has been established to be fit for its intended use in measurement or in examination of nominal properties.
2. **Measurement procedure** – a description of a measurement according to one or more measurement principles and to a given measurement method, based on a measurement model and including any calculation to obtain a measurement result. A measurement procedure is usually documented in sufficient detail to enable an operator to perform a measurement.
3. **Measurement unit** – a real scalar quantity, defined and applied, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a unit.
4. **Combination** – a combination of the standards.

3.2 Entities in a conceptually modeled measure

NOTE: *Different disciplines have different measurement ontologies.*

An entity is a conceptual categorization of information defined by common attributes and constraints at the systems level. The term corresponds to the "Entity" construct as defined in ISO 10303-11.

There are at least two principal types (classifications/categorizations) of entities in measurement [systems]: objects (i.e., events, values, methods) and properties (i.e., attributes and characteristics:

Objects (and events, respondents) - Objects are fundamental blocks of understanding (as in, unitizations of information). Objects are the entities ("building blocks") that compose a mental model of the world. Every "thing" is an object (as a significant, bounded information patterning set). Objects are, for example, phenomena, bodies, or substances, but also individuals, processes, and organizations. Objects and events are described through constraints

1. **Events** – An sensed or experienced interaction [between two or more differentiated objects].
2. **Methods** - A method is an action that an object can perform. An algorithm is a pre-set method. A method is a formal ordering of constraints.
3. **Numerical values (numbers, a syntactic category)** – A unit of information signifying an ordered rank meaning that expresses the magnitude (amount or quantity) of a fundamental iterating pattern. In measurement a number is not usually used by itself, but in tandem with some other term, its dimension[al property], which will normally correspond to some Aristotelian category of substance or quality. Numerical values are assigned to properties as the result of measurement. In application, a number is a relation between the concepts of 'magnitude' (amount) and 'unit' (reference).
4. **Properties (and attributes, characteristics)** – A 'property tag' names what is being measured. This is the measured property or attribute, which is organized into a relational system otherwise known as a measurement classification (or taxanomical, ontological) system. Herein, a property is either:
   A. A quality (characteristic or attribute) of an object (or event), or
   B. An aspect of its behavior.

Note that it is sometimes said that attributes are properties of entities. In this sense, there are two principal categories of attributes:

1. Internal attributes (direct measures) are measured directly from the entity.
2. External attributes (indirect measurement) are indirectly measured.

Properties are, for example:

- Length, loudness, and frequency.

The ISO VIM3 states that a 'property' is either a nominal property or a quantity, and a quantity is either an ordinal quantity or a Euclidean quantity for which a unit can be defined. Hence, a property is one of the following:

1. A **nominal property or qualitative property (no magnitude or scale)** - Nominal properties cannot form scales. Expressed by categories (names) in a set. A nominal property is a property

of a phenomenon, body, or substance, where the property has no magnitude. A nominal property is a property that cannot be ordered according to magnitude. For example, the sex of a human being cannot be ordered according to magnitude – in normal physiological procreation, two sexes are required (i.e., there is no magnitude between each other, or the top-level category). In some disciplines, the term qualitative analysis is used to describe the examination of nominal properties. Qualitative analysis produces [measurement] data acquired without magnitude. In the previous sentence, the term measurement is crossed out; this is because, it is possible to measure a quantity, whereas obtaining information about a nominal property is not a measurement.

2. A **quantity [property] (measurable property)** - Quantities form scales, continuums. If it is not a nominal type property, then it is a quantity type property. A quantity is any property that has a size (magnitude) that can be evaluated (compared and integrated) through some measurement. 'Quantity' is a specific type of property. Only quantities [of phenomenological objects or events] are technically measurable. Each sub-type of this property has an accompanying application as something called a [quantity] 'scale'. A quantity scale (a.k.a., measurement scale) is an ordered set of values of quantities of a given kind used in [sequential] ranking, according to [the order of] their magnitude. Types of properties (e.g., nominal, ordinal, quantity, cardinal) become scales of quantities (e.g., nominal, ordinal, interval, ratio).
    A. An **ordinal quantity** [property] scale - expressed by ordering of categories in a set.
    B. A **cardinal quantity** [property] scales also known as a Euclidean quantity [property] which must have defined units. Expressed by a number and a measurement unit as part of a system of [existent] quantities. The physical quantities of the universe are cardinal/Euclidean.

Geometric quantities are paradigmatic of measurable entities. Hence, quantity is an axiom of measurement – quantity grounds the theory of measurement. Because measurement requires quantity, and quantity is (axiomatically) logically numerical, the foundations of measurement can be notated in purely mathematical terms.

NOTE: *The division of the concept of 'quantity' according to 'kind of quantity' is arbitrary to the extent that the unified principles of the universe are not yet known.*

The cardinality and ordinality:

1. Cardinal has to do with cardinality or the magnitude or quantity of things.
2. Ordinal has to do with ordinality or the ordering or ranking of things.
3. Thus, first is an ordinal number. Its cardinal equivalent is one.

3.3 Ordinal quantity (ordinal property)

An **ordinal quantity** is a quantity defined by a conventional measurement procedure, for which a total ordering relation can be established, according to magnitude, with other quantities of the same kind, but for which no algebraic operations among those quantities exist. Ordinal quantities are usually not considered to be part of a system of quantities, because they are related to other quantities through empirical relations only.

Examples of ordinal quantities (in applied scale form) are: Rockwell C hardness scale, Octane number for petroleum, and the strength of an earthquake on a Richter scale. The numbers on these scales are arbitrary and dimensionless.

Ordinal quantities have neither measurement units nor quantity dimensions. Ordinal quantities are arranged according to ordinal quantity scales.

3.3.1 Ordinal quantity scale (ordinal scale)

An ordinal quantity scale (ordinal scale) is a conventional reference scale or a quantity scale, defined by cooperation, on which only comparison of magnitude applies. An ordinal quantity scale may be established by measurements according to a given measurement procedure. Also of note, ordinal quantities are ordered on ordinal quantity scales.

3.4 Combining entities in measurement

Measurement combines the categories of quality and quantity in order to establish the quantity of a particular quality. Quantities are fusions of numbers and property dimensions (in metrology, the latter are called 'quantity dimensions', though there also exist qualified properties).

'Entity value' principle states that no entity can possibly at one and the same time take two specific values of the same property dimension (quantity variable). For example, no material object can simultaneously have two masses, two volumes, two electric charges, etc.

The concept of dimension is axiomatic to [material] existence, representing the class of information about which reality itself is composed. The spatial, material system is conceived to have the following initial dimension: length (x-axis), width (z x-axis), height (y x-axis), and time (technically, "space-time/memory"). Here, a 'dimension' is a "pure" measurement, as opposed to a scale, which is a ratio of measurements (e.g., kilometers per hour or amps per second, versus mass or temperature).

The principal [visualization] tool in measurement is 'scale'. A scale is a visualization tool that precisely enables deduction of a value of a given quantity (magnitude or amount) by knowing its position [on the scale] and the scale's ratio [between one position and the next]. A scale is the standard (reference) and scope (boundary) of measurement (e.g., nominal, ordinal, ratio scale, etc.).

A self-organizing system can encode the concepts of objects and properties (relationships) to form scales (visual expressions with position and ratio information), upon which logical processing (i.e., mathematical operations) may be performed.

Here, the concept of a [measurement] unit provides meaning to the 'scale' by differentiated identification of one categorical unit from other axiomatic or derived units).

For the supra-system, the level of measurement (i.e., scale of measurement, property scale, or variable scale) determines how the data will be interpreted (i.e., what mathematical operations can be performed). Therein, knowing the level of measurement resolves what statistical analysis is appropriate on the values that were assigned [to the variable at that level].

When there is a scale of possibilities, there becomes a need for defining quantifiable measures for the optimal functionality of a system. That optimally functional or desired value, of a measured system, is called a 'metric'.

A scale is required in measurement for a specific value, among a sequence of possible values, to have meaning.

3.5 Measurement scales

Measurement scales are the symbolic representation of possible measurement results. Measurement scales are used to categorize and/or quantify variables so that correct mathematical operations may be applied. Each additional mathematical operation generates a new "scale of" measurement. In general, there are four scales of measurement:

1. There are four scales of measurement commonly used in statistical analysis.
2. There are four types of data commonly used in information processing.
3. There are four types of variables commonly used in quantifying and qualifying.

Those four categories (i.e., scales of measurement) are, in order:

- Nominal, ordinal, interval, and ratio.

The scientifically accepted physical quantity-value scaled units are (Read: the fundamental/base quantity values are):

- ...
- The meter scale – property is length units.
- The kilogram scale – property is mass units.
- The second scale – property is time units.
- ...

3.6 Variables in measurement

> **NOTE:** *Measurement operationalization is the process of developing specific variables that will be used to measure a concept.*

A variable is any entity that can take on different values. In statistics, where variables are actually used, a variable is any characteristics, number, or quantity that can be measured or counted. A variable may also be called a data item.

Further, in statistics, the general property that is being measured through one or another of the three fundamental measurement processes (counting, ordering, sorting) is termed a 'variable'. Any particular measured instance of that property is spoken of as a 'variate'. 'Variate' is a single variable instance.

The term 'variable' implies that the results of the measurement process are capable of varying from one time to another or from one item to another. For instance, the categorical measurement of gender among a mixed group of human subjects will vary from one subject to another between the two possible outcomes, female and male.

A specific variable represents a specific concept with a logical indicator or value. It is a data point that can be counted, ordered, or sorted. Strictly speaking, measurement does not occur on things, or qualities, or properties, but "indicants" of properties.

The opposite of a variable is a constant. A constant does not vary from one time to another or from one item to another. It is an unchanging value that will applied mathematically to a data set.

In measurement, the word remains with a common meaning, but is often used in multiple different contexts. The following are the multiple ways in which the term 'variable' may be applied.

In measurement, a variable is:

1. A [measurable] property (attribute or characteristic) of an object or event (of existence) that can be assigned a number (numerical variable) or a category (categorical variables), and
2. Is expected to change over time (measurement variables).

In measurement, there are two types (categories) of variable [processes]:

1. **Qualitative variables (categorical variables)** - A qualitative variable is one which measurement occurs with categories possessing no meaningful numerical values.
2. **Quantitative variables (measurement variables)**

- A quantitative variable is one which measurement occurs with meaningful numerical values.

There are different ways variables can be described according to the ways they can be studied, measured, and presented. In common application, entities become types of variables [in numerical-mathematical scales of operation], whereupon variables are typically classified as either of two types:

1. **Categorical variables (a.k.a., qualitative variables)** - Categorical variables are variables whose levels are distinguished by name only. Properties become categorical variables. Categorical variables have values that describe a 'quality' or 'characteristic' of a data unit, like 'what type' or 'which category'. Categorical variables fall into mutually exclusive (in one category or in another) and exhaustive (include all possible options) categories. Therefore, categorical variables are qualitative variables and tend to be represented by a non-numeric value.
   - Categorical variables may be measured on one "scale": nominal.
2. **Numeric variables (a.k.a., quantitative variables or measurement variables)** - Numerical values become numerical variables. Numerical variables have values designated by numbers that have some meaning relative to one another. Numeric variables have values that describe a measurable quantity as a number, like 'how many' or 'how much'. Therefore numeric variables are quantitative variables.
   - Numerical variables may be measured on three scales: ordinal; interval; and ratio.

Numeric, quantitative measurable variables may be further described as either continuous or discrete:

1. **Continuous variables** – variables that have an infinite (or significantly large) number of possible values. A continuous variable is a numeric variable. Observations can take any value between a certain set of real numbers. The value given to an observation for a continuous variable can include values as small as the instrument of measurement allows. Examples of continuous variables include: height, time, age, and temperature.
2. **Discrete (meristic) variables** – variables that only have whole number values. A discrete variable is a numeric variable. Observations can take a value based on a count from a set of distinct whole values. A discrete variable cannot take the value of a fraction between one value and the next closest value. Examples of discrete variables include the number of registered cars, number of business locations, and number of children in a family, all of which measured as whole units (i.e. 1, 2, 3 objects).

The data collected for a numeric variable are quantitative data.

Categorical, qualitative variables may be further described as:

1. **Nominal variable** - a categorical variable. Observations can take a value that is not able to be organized in a logical sequence. Examples of nominal categorical variables include sex, business type, eye color, religion and brand.
   - **Qualitative-nominal** – qualitative variables where the categories have no natural ordering.
2. **Ordinal variable** - a categorical variable. Observations can take a value that can be logically ordered or ranked. The categories associated with ordinal variables can be ranked higher or lower than another, but do not necessarily establish a numeric difference between each category. Examples of ordinal categorical variables include academic grades (i.e. A, B, C), clothing size (i.e. small, medium, large, extra large) and attitudes (i.e. strongly agree, agree, disagree, strongly disagree).
   - **Qualitative-ordinal** – qualitative variables where the categories have a natural ordering.
3. **Qualitative-dichotomous** – qualitative variables with two categories.

The data collected for a categorical variable are qualitative data.

CLARIFICATION: *The words "measurement variable" are used here in reference to two related things. First, the term 'measurement variables' refers to all possible variables in measurement (as a concept, quantity and quality variables), and secondly, the term 'measurement variable' refers to only quantitative variables.*

In experimentation and measurement data acquisition, there are two axiomatic (principal or ontological) categories, each with two principal types of variables:

1. **Categorical variables** – a variable that can be placed into categories, but these categories may not have any logical ordering. A categorical variable is a property of an object which can be broken down into different classes or categories.
   A. *[Scale level 1]* **Nominal variables** – classification is made into unordered categories. Nominal variables are expressed as names (such as "female"). Nominal variables classify

observations into discrete[ly named] categories.

B. *[Scale level 2]* **Ordinal variables (ranked variables)** – classification is rank ordered on some characteristic. However, there is no indication of how much greater one instance is than another. These are expressed as positions (such as "third"). Ranked variables, also called ordinal variables, are those for which the individual observations can be put in order from smallest to largest, even though the exact values are unknown.

2. **Measurement variables** (a.k.a., numeric variables or quantitative variables) – a measurement variable is one where numerical values can be assigned and objects or events can be ordered according to those values. Measurement variables are expressed as numbers and a reference (such as 3.7 mm).
   A. *[Scale level 3]* **Interval variables** – values for interval variables have equal intervals between them; however, they lack an absolute zero point.
   B. *[Scale level 4]* **Ratio variables** – values for ratio variables have equal intervals between values, and there is an absolute zero point.

The principal [measurement] variable from which all other variables (except nominal) are derived is 'quantity'. Quantity is the source conception of a 'measurement variable' -- if there is a potential differentiation for that which may be known to exist, then what is the separation?

> **NOTE:** *The mathematical theories underlying statistical tests involving measurement variables assume that the variables are continuous. However, [continuous] statistical tests also work on discrete measurement variables. The only exception is when there is a small number of possible values of a discrete variable, in which the variable may be treated as nominal (instead of, a measurement variable).*

In the application of statistics to measurement variables, there is the possibility of calculating for more than one numeric value for a variable:

1. A measurement variable with only two values should be treated as a nominal variable;
2. A measurement variable with six or more values should be treated as a measurement variable;
3. A measurement variable with three, four, or five values requires complex simulation.

3.6.1 Numeric variable scales

1. **The cardinal number scale and cardinal measurement** - In the cardinal measurement there are two subcategories, ratio scale and interval scale, and all cardinal variables are either continuous or discrete.
   A. **Discrete cardinal variables** – count variables. For example, number of people in a town, family size, number of books, number of heads in 10 tosses of a coin, and so on. Discrete variables can have negative values; for example, if the net change in demand is measured by the difference of arrival of customers the result can be negative or positive. Discrete numerical variables are variables that can take on only whole number values. Discrete numerical variables are typically the result of the counting operation/process (e.g., counting things, events, activities, types).
   B. **Discrete scale of measurement** - Discrete cardinal variables - Discrete/Integer scale of indivisible units: 1,2,3,4,5,6,...
   C. **Continuous cardinal variables** - All these cardinal variables (time, height, weight, distance) are examples of Continuous variables - they are measured in real numbers and they have unit of measurement. Continuous numerical variables are variables that can take on any value whatsoever. They can be whole numbers, or they can be numbers to any number of radix points (e.g., decimal points - fractions of a whole number).
   D. **Continuous scale of measurement** - Continuous/ratio[nal]/fraction scale of [in principle] infinitely divisible units: 1.23,2.9120,4.323442,...
2. **Equal interval scale** – equal intervals exist between their successive units of measurement. If a measurement scale possesses this property, then it is possible and meaningful to take two or more measures from that scale and perform the simple arithmetic operations of addition and subtraction.
   - **Ratio scale** – a point is designated as zero, which represents an absolute zero of the quantity that is being measured (e.g., zero length represents the absolute absence of length). Scales of measurement that have both equal intervals and absolute zero points are known as ratio scales.
   - **Non-ratio scale (interval scale)** – a point [on the scale] is either:
   - Not designated as absolute zero of the quantity that is being measured (e.g., kelvin temperature scale).
   - Or, the designation of zero is only an arbitrary point that happens to be called "zero" (e.g., celsius temperature scale).

3.7 Conceptual mapping of the empiric, real world through qualification and

quantification

Conceptual mapping of the empiric, real world is carried out through [at least] two processes, qualification and quantification.

Note here that the terms 'qualification' and 'quantification' both end with the suffix "-fication". The suffix-noun "-fication" means - making, producing, or representing. For example, reification means to making something real or physical (such as, making a clay pot). However, the concept 'reification' can also be applied philosophically. For example, an actual 'shadow' is the absence of light, where light is an actual thing. The shadow is not the presence of a thing, but its absence. A shadow cannot be reified; it cannot be experienced and conceived of as a separate object/thing. To make the experience of an absence [of a thing] into a thing itself is bound to cause instability in a societal trajectory toward fulfillment and ecological well-being. It could be said that qualification and quantification depend to a large degree on accurately experiencing, and hence conceiving, of the real world. When absences are turned into qualified things and then quantitatively measured, that data may still have usefulness, but the context in which it was

Human cognition can recognize patterns of quality and quantity in our environment. These patterns are mapped to concepts. In concern to measurement, the supra-mapping conceptions are 'quality' and 'quantity'. Whereas quality represents categories and their ordering, quantity represents the presence of a meaningful number.

The properties, characterizations, and attributes of existence can be categorically described in two ways: qualitatively (through words and linguistics) and quantitatively (through numbers and mathematics).

There are two principal descriptive forms (notations or expressions) of 'measurement', in the most general use of the term:

1. A **qualitative description** is the use of words and linguistics.
2. A **quantitative descriptio**n is the use of numbers and mathematics.

Hence, there are two principal types of measurement:

1. **Qualitative measurement** uses words, representing linguistic, semantic conceptions, to describe [that which is/was existent] in relation to a model (scheme) conceived to pre-exist. Qualitative measurement requires the assigning of a word (concept) capable of functioning under linguistical logic (i.e., in a linguistic system).
2. **Quantitative measurement** uses numbers, representing mathematical conceptions, to describe [that which is/was existent]. Quantitative measurement requires the assigning of a number (value or count) capable of functioning under mathematical logic (i.e., in a mathematical system).

Measurement involves processes that determine the value of a [new] quantity or quality of [some category of] information.

1. **Qualitative information** - involves processes that determine the value of [new] qualitative [information].
2. **Quantitative information** - involves processes that determine the value of a [new] quantity [of information].

In this sense, there are two general types of measurement data (and research):

1. **Qualitative data (qualitative research)** is information about qualities; information that cannot be expressed or processed through numerical conception. Qualitative data involves linguistic characteristics and descriptors that can't be measured, but can be observed subjectively.
2. **Quantitative data (quantitative research)** is information about quantities; that is, information that can be measured and written down with numbers. Quantitative data involves numbers and systems that can be measured objectively. When something [existent] is "measured", then the result is quantitative data. All [numerical] measurement is quantitative data.

And, in measurement experimentation there are two types of variables (Read: a concept, factor, trait, condition, behavior, etc) of an object or system (in the real world) that can exist with differentiation (i.e., in differing amounts or types):

1. **Qualitative variables** - take on values that are names or labels.
2. **Quantitative variables** – take on values that are numeric.

3.8 Mapping process categories

There are two mapping process categories for mapping existence to workable information sets. The two process categories are:

1. **Qualify[ing]** means to characterize by naming an attribute; it means to state any property or characteristic of something. Qualify refers to meeting the terms of eligibility or criteria.
   - There is categorization by descriptive values (categorical values).
2. **Quantify[ing]** means to find, determine, or otherwise calculate the quantity or amount of (something). In application, quantify is describing

[some thing] numerically.
- There is categorization by numbers (numerical values).

Measurement is the process of assigning to some specific instance (of a category of existence), a numerical value (quantification) and/or qualifying condition (qualification). All quantities (quantity values) are actually qualified by [their] units, which represent either a qualifying procedure and/or a qualified definition.

> **NOTE:** *"Qualify" is also defined in common parlance as: to have the necessary skill, knowledge, or other requirements to do a particular process, activity, or to have the qualifications to do something.*

3.9 Quantity defined by standards

ISO 80000-1:2009, 3.1: The International Standards Body defines a 'quantity' as a property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed by means of a number and a reference.

A reference can be:

1. A measurement unit.
2. A measurement procedure.
3. A standardized reference material/tool
4. Or, a combination of these.

Simply, a quantity is a property of a phenomenon, body or substance, to which a number can be assigned with respect to a reference (of which there are four possible reference types).

Quantity is a specific type of property. And, only quantities [of phenomenological objects or events] are technically measurable.

International Vocabulary of Metrology 3rd edition (VIM3): starts with a definition of 'quantity' (def. 1.1) followed (1.2) by one for 'kind-of-quantity'. Two other VIM3 definitions relevant are those of 'quantity dimension' (1.7) and 'quantity value' (1.19).

- 1.1 quantity = property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference.
- 1.2 kind-of-quantity = aspect common to mutually comparable quantities.

> **Insight:** *The first axiom of measurement is quantity, and the second, uncertainty.*

VIM3 then defines quantity value as an expression that is related to a spatio-temporally localized (individual) quantity (i.e., a quantity value is a representation of a (individual) quantity). The definition says:

- 1.19 quantity value = number and reference together expressing magnitude of a (individual) quantity [1, p.12, the parenthesis added].
- 1.7 quantity dimension = expression of the dependence of a quantity on the base quantities of a system of quantities.

Continuous quantities possess a particular structure that can be explicitly characterized as a set of axioms that define such features as identities and relations between magnitudes (sequences of patterns). In science, **quantitative structure** is the subject of empirical investigation and cannot be assumed to exist a priori for any given property.

Every quantity structure has the following fundamental characteristics:

1. Relationships of equality or inequality can in principle be stated in comparisons between particular magnitudes, unlike quality, which is marked by likeness, similarity and difference, differentiation.
2. Additivity may involve concatenation, such as adding two lengths A and B to obtain a third A + B. Additivity is not, however, restricted to extensive quantities but may also entail relations between magnitudes that can be established through experiments that permit tests of hypothesized observable manifestations of the additive relations of magnitudes.
3. Continuity, as a type of quantitative attribute, where continuity means is that if any arbitrary length (dimension), a, is selected as a unit, then for every positive real number, r, there is a length b such that b = ra.

3.9.1 Qualifiers

Quantifiers are words and phrases used to indicate quantity. These include, but are not limited to:

1. A number
2. Few
3. Many
4. Each / every
5. Several
6. An amount
7. Little
8. Less
9. More
10. Much
11. All
12. Some

3.10 Quantity commonly defined

Take note that synonyms for quantity include:

1. Sequence
2. Magnitude
3. Amount
4. Size
5. Degree
6. Weight (not the tool, 'scale')

A quantity is some measured or measurable amount (i.e., quantity or sensation) of some "thing" (of a pre-existing pattern). Therein, a 'unit' of measurement is assigned to selectively identify and categorize (tag, name) the concept[ual thing or dimension] being measured. A quantity is a quantifiable numerical assignment of some property, which is conceptualized as a particular phenomenon (natural process), body (object), or substance (material).

CLARIFICATION: *Magnitude (size) means the numerical value which tells the amount of that physical quantity.*

In measurement, the terms quantity, quantity value, and value, can mean the same thing:

1. A 'quantity' is an amount of something that must have a value.
2. A 'quantity value' is a number and reference together expressing the referential magnitude of a quantity.
3. A 'value' is a number with a reference.

Terminological clarification:

1. **Quantity kind or type (quantity dimension)** – any observable property of any object that can be measured and quantified numerically. A quantity is any property which has size (magnitude) that can be evaluated (compared and integrated into an information model) through some measurement process.
   - For example: length, mass, time, force, energy, electric charge.
2. **Quantity** – observable property of a particular object that can be measured and quantified numerically.
   - For example: length, mass, speed, temperature of a particular object.
3. **Quantity value** – Magnitude of a quantity expressed as a product of a number and a unit.
   - For example: a velocity of m/s.
   - The term 'indication' (result) is used to express the quantity value provided by a measuring instrument.

INSIGHT: *A quantity is anything that can be measured.*

In practice, the terms 'dimension' and 'quantity' tend to become synonymous. Each base quantity is regarded as having its own dimension, and the dimension of a derived quantity is contains the same information about its relation to the base quantities as that provided by the SI unit of the derived quantity as a product of powers of the SI base units. A quantity is also sometimes called a 'quantified dimension'.

Table 39. Measurement > Quantity Comparison*: Table shows two examples (length and power) of the physical dimensions of quantity, physical dimension, and unit.*

Quantity	Dimension	Unit
Length	L	Metre
Power	ML^2T^{-3}	Js^{-1} or watt

Relations between different quantity types/ dimensions are defined by units. A unit is a particular physical quantity, defined and adopted by convention, with which other particular quantities of the same kind are compared to express their value. All physical quantities can be expressed in terms of seven base units.

3.11 A system of quantities

All quantities together with their defined relations form a 'quantity set', otherwise known as a 'system of quantities'. A system of [physical] quantities is a set of quantities together with a set of non-contradictory equations relating those quantities:

1. Quantity objects (quantities) - Base/fundamental and derived quantities.
2. Relational objects (equations) - a set of non-contradictory equations relating those quantities.

In order to establish a system of units, such as the International System of Units (SI), it is necessary first to establish a system of quantities, including a set of equations defining the relations between those quantities.

That which indicates the thing being measured is the reference part, and that which indicates the numerical result of the measurement is the number part:

1. The number part is called a numerical value. Take note here that the number part is also sometimes referred to by just the word "value".
2. The reference part is an entity called a measurement unit, which is defined (VIM, Section 1.9) as a "real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number".

NOTE: *Quantity values are viewed as data by a supra-system.*

3.12 Expressing quantity (in natural language)

A quantity is the combination of a [sequenced] number and a [referential] unit, where the unit may be "pure" (base/fundamental), or a ratio of two relatable units (that describe some function present in the real physical world).

Quantity is expressed by three principal elements (of naming):

1. Identifiers (definite and indefinite) – identify a class of object ("thing") or an example of a class.
   A. **Indefinite identifiers** – indefinite articles (a or an) and the zero article.
   B. **Definite identifiers** – definite article (the), some pronouns (the demonstratives, possessives, anaphoric pronouns), and ordinal numbers. Note, the definite article "the" identifies the particular class.
2. **Quantifiers (definite and indefinite)** – express (or otherwise specifically identify) a quantity.
   A. **Indefinite quantifiers** – express uncertain quantities, such as: a few or a little; many or much; a great number or a great amount of; several; all; plenty of; a lot of; enough; some; any; every; no. The most common logical quantifiers are:
   B. "All" refers to the whole number, to all the elements or units.
   C. "Some" refers to a portion of the whole number, elements, or units.
   D. "Every" refers to all taken separately.
   E. "Each" refers to one example of a class of all.
   F. "Any" refers to all classes.
   G. "None" refers to no one of all.
   H. Definite quantifiers – are cardinal numbers.
3. Nouns of the following three types:
   A. Count unit nouns or countables.
   B. Mass nouns, uncountables, referring to the indefinite, unidentified amounts;
   C. Nouns of multitude (collective nouns).

The word 'number' is a noun of multitude, and stands for either:

1. A single entity, or
2. the individuals making the whole.

3.13 Classifying property-quantities: system quantity dependency

> **NOTE:** *The type of quantity to be measured also impacts measurement requirements.*

Some quantifiable properties (i.e., quantities) are dependent on the quantity, size and extent of the system of which they are a part; and, others are not dependent on the system's quantity.

A quantity, as a sub-part of a system, can depend, or not depend, on the size of the system itself or the quantity (amount) of matter (mass) present in it. A quantity that does not depend on (i.e., is independent) the size of the system expresses an in-tensive type of property/quantity, and a quantity that does depend on the size or extent of the system expresses an ex-tensive property.

> **NOTE:** *The term, tensive, means causing or expressing tension.*

In physics, a fundamental distinction is made in measurement between intensive and extensive quantities (here, a quantity is a property):

1. **Intensive property** - The magnitude of an intensive quantity is independent of the mass of its system (e.g., temperature, density, or pressure). Here, the word quantity may be replaced with property. As in, an intensive property is a physical property of a system that does not depend on the system size or the amount of material in the system. They are independent of the quantity of the system; it is independent of mass. They are independent of the size or extent of the system.
2. **Extensive property** - The magnitude of an extensive quantity is additive (like mass, volume, or energy). Here, the word quantity may be replaced with property. An extensive property is additive for subsystems. This means the system could be divided into any number of subsystems, and the extensive property measured for each subsystem; the value of the property for the system would be the sum of the property for each subsystem. They depend on the quantity of the system; it is dependent on mass (as a variable). They depend on the size and extent of the system.
   A. Extensive quantity
   B. Potential energy (non-conserved quantity):
   C. Unavailable energy;
   D. Available energy (heat interaction as entropy, and work interaction)
   E. Mass (conserved quantity)

Note that in some disciplines, there is no recognized type distinction between intensive and extensive quantities.

3.13.1 Quantity and quantity value in mathematics

> **NOTE:** *Mathematically, a quantity is a scalar. However, a vector or a tensor, the components*

of which are quantities, is also considered to be a quantity.

From a mathematical perspective, a 'quantity value' is an algebraic term. In algebra, the concept 'term' represents a mathematical expression composed of two different parts: the number part (a.k.a., numerical coefficient) and the variable part (often notated as "x" or "y"). Similarly, the result of a measurement is, mathematically speaking, a 'term', for there is a number part and a unit (variable) part, which may be expressed [in notation] with constants and other variables, and equality symbols may be included to form equations from which [statistical] mathematical operations may be performed [to acquire more/new information].

The language of algebra has no meaning in and of itself. The theoretical mathematician deals entirely within the realm of the formal language and is concerned with the structure and relationships within the language. The applied mathematician or statistician, on the other hand, is concerned not only with the language, but the relationship of the symbols in the language to real world objects and events. The concern about the meaning of mathematical symbols (numbers) is a concern about measurement.

Magnitude (how much?) and multitude (how many?) are the two principal types of quantities, which may be further divided as mathematical and physical. The essential part of mathematical quantities consists of having a collection of variables, each assuming a set of values. These can be a set of a single quantity, referred to as a scalar when represented by real numbers, or have multiple quantities as do vectors and tensors, two kinds of geometric objects.

Quantities can be used as arguments of a function, variables in an expression (independent or dependent), or probabilistic as in random and stochastic quantities.

Number theory covers the topics of the discrete quantities as numbers: number systems with their kinds and relations. Geometry studies the issues of spatial magnitudes: straight lines, curved lines, surfaces and solids, all with their respective measurements and relationships.

Algebra operations are used for performing computations with quantities. Here, algebra operations allow for computations with uncertain values. These operations enable model-level simulations that consider data uncertainty and units (encoded through referential databases). In application, [physical] information processing requires a computational kernel for computing quantities.

NOTE: *Quantities may be integrated with modeling language, as in the case of UML.*

3.14 Physical and non-physical quantities

INSIGHT: *All physical measurements are geometric measurements.*

There are two categories of quantities as viewed from the physical, material perspective:

1. A **physical quantity** is a quantity that can be used in the mathematical equations of science and technology. Systems exist along a spatial-temporal continuum. Physical quantities are used in science and engineering because they are objective, and hence, may be used for logical inquiry and construction.
   A. The material properties of the surrounding world include: The existence, operation, placement, and composition of material (spatial-temporal) objects.
   B. Sud-divided into base and derived quantities.
2. A **non-physical quantity** is a quantity that cannot be measured by any mean or media. These quantities do not have magnitude of themselves. A non-physical quantity is a qualitative measure (and, non-physical quantities can have order). Some common example of non-physical quantities are: feelings, angriness, rudeness, etc. For these measurements (as in, the measurement of non-physical quantities), it is not possible to ensure traceability because of their exclusive nature as substantiated solely by the author. It is relevant to note here that it is scientifically understood (in neurophysiological flow literature) that biology responds before psychology (or with psychology, in the case of highly intelligent consciousness). And, all aspects of biology and electromagnetism area measured as physical quantities (and not, non-physical quantities). Biology can be quantifiably measured.
   A. The immaterial properties of the inner world include: anything experiential which can be described as feeling or e-motion. Non-physical quantities exist only in the mind of people either as reflections of properties of the real world, or in the form of people's own understandings (...to which the body responds faster than the mind, and can be accurately, traceably quantified).
   B. Immaterial properties do not only exist in people's minds without any material dimension, in fact, the human body express in more or less optimal, or disturbed, functioning.
   C. Sud-classified (sub-divided) into simple (a.k.a., base) and compound (a.k.a., derived) quantities. The simple physical quantities are constituent parts of the compound quantities, the latter being composed of a set of simple or compound quantities of a lower level.

INSIGHT: *The human organism cognates the* *material world by means of sense organs, through organoleptic measurements of its quantitative characteristics, while the immaterial world is perceived through measurements of its qualitative characteristics.*

When measuring physical quantities, standard reference objects and/or machines may be used, such as: length gauges (rulers) to measure length, and mass gauges (weights) to measure mass.

It is not accurate to say that non-physical quantities (qualities) only measure human opinion; they can also measure felt (experiential) fulfillment. Organisms express bio-electrical responses to particular environments. Therein, there is an ordering between suffering and well-being.

There are cases where the non-physical (immaterial) quantity being measured is a human opinion. Human opinion is subjective estimation. And, human opinion is largely dependent on an individual's specific life circumstances (social, decision, lifestyle, and material, without reference to an exist world). Opinion is largely dependent on exposure to information, personal preference, social influence, personal well-being, environmental factors, tastes, health, etc.

Here, the measurement of an immaterial property (opinion) comes down to a comparison of manifestations of this property, and, as a result, the question about when the property has manifested itself in a greater degree can be answered.

From the acquisition of opinions expressed about the manifestation of an immaterial property comes a comparative, quantitative estimate (i.e., a single measurement is taken on an ordinal scale). Here, one opinion is one measurement unit (along an ordinal [quantity] scale). And, the number of opinions is a quantitative characteristic of a non-physical quantity (a quality).

Note that besides felt human fulfillment and ecological well-being, human opinion has no fixed dimension, and hence, without the concepts of fulfillment and well-being, it is separated from a unified model of understanding and constructing.

3.15 Classifying physical quantities

Physical quantities can be classified in a number of ways:

- **Electrical quantities:** resistance, capacitance, permeability, permittivity (voltage, current, inductance, electrical power, electrical energy).
- **Non-electrical quantities:** fluid pressure, displacement, torque, temperature, area, volume.

The concept of a physical quantity can be classified according to whether it is electrical or non-electrical:

- **Length-type quantities:** The quantities diameter, circumference, and wavelength are generally considered to be quantities of the same kind, namely of the kind of quantity called length.
- **Energy-type quantities:** The quantities heat, kinetic energy, and potential energy are generally considered to be quantities of the same kind, namely of the kind of quantity called energy.

Note: *Many traditional economists hold the view that utility is measured quantitatively, like length, height, weight, temperature, etc. This concept is known as cardinal utility concept. On the other hand, ordinal utility concept expresses the utility of a commodity in terms of 'less than' or 'more than'.*

3.15.1 The dimensional property attribute of [classified] physical quantities.

A quantity as an information data point may be either a scalar quantity or a vector quantity. In physics, there are two principle types of physical quantities [that can be measured]: scalar quantities and vector quantities. These two categories are typified by what information they require. Scalars require one piece of information (a number), and vectors require two pieces of information (a number and [coordinated] direction). A scalar measurement is the measure[d result] of a scalar quantity, and a vector measurement is the measure[d result] of a vector quantity.

3.15.2 Scalar – magnitude only (a scalar represents the magnitude or size of a quantity)

A scalar variable is a variable that holds an individual value (single number). A scalar number is a number used to measure some quantity to any desired degree of accuracy.

1. **Scalar quantities** - Scalars are used to describe one dimensional quantities, that is, quantities which require only one number to completely describe them. A scalar quantity represents a physical quantity specified by magnitude. Scalar [quantities] are physical [quantities] represented by a single number [magnitude] and no direction. Scalars can be represented by a |x| matrix. In visualization, scalars are numbers.
    A. Examples include, but are not limited to: temperature, time, height, speed, mass, volume, location along a line (1D). Position and distance are scalars, because there is no direction.
    B. one scalar • another scalar = a scalar
    C. one scalar • a vector = a vector
2. **Scalar measurement** – a numerical descriptive signifier of a quantity, magnitude, or size of a bounded sensation. A scalar measurement can be

represented with a number alone (with relevant units). It describes a quantity, magnitude or size of a measurement alone. For example, mass and temperature are scalar measurements.

3.15.3 Vector – magnitude and direction (a vector represents the magnitude, size and direction of a quantity)

A vector variable is a variable that holds more than one individual value.

1. **Vector quantities** - Vectors are used to describe multi-dimensional quantities. Multi-dimensional quantities are those which require more than one number to completely describe them. Vectors, unlike scalars, have two characteristics, magnitude and [a systematically coordinated] direction. The magnitude of a vector is its “length” (or other quantity in some units). Vector quantities are [not necessarily physical quantities and] are represented by a number (magnitude) and a direction. The direction is usually given in terms of some angle. Vectors can be represented by a lxA ([4 2], row vector) or an nxl ([4 / 2], column vector) matrix. Vectors are one dimensional. In visualization, vectors are arrows.
    - Examples include, but are not limited to: location in a plane (2D), location in space (3D), velocity, acceleration, force, displacement, momentum.
    - one vector • another vector = a vector
2. **Vector measurements** – a numerically descriptive signifier of the relationship between the two fundamental dimensions of magnitude (size or quantity; inertia and acceleration) and direction (force and motion). Vectors are a form of measurement that conveys both magnitude (size or quantity) and direction (with relevant units). Velocity is a good example of a vector measurement (the object moves at 3m/s to the East). It is not to be confused with speed, which is scalar (e.g., the object moves at 3m/s). When visualizing vector measurements, an arrow is a common symbol for the vector.

Vectors (vector numbers) can be added together in ways that scalars (scalar numbers) cannot.

> **NOTE:** *A vector space is defined as a set of vectors, a set of scalars, and a scalar multiplication operation that takes a scalar k and a vector v to another vector kv.*

The term “scalar” comes from linear algebra, where it is used to differentiate a single number from a vector or matrix.

3.15.4 Standard scalar measurement

In standard scalar measurement, points exist along a principal standard measurement scale, the scale of cardinal numbers. The following are ways of taking a standard scalar measurement; if, for example, “you” measure the width of “your” desk, “you” take a tape measure and align it with a point on the desk, and then, “you” count off the number of centimeters or inches. Or, if “you” measure the outdoor temperature at the present moment, “you” take a thermometer outdoors and count off the number of degrees Celsius or degrees Kelvin. If “you” are sitting in a room measure the number of humans in the room, “you” count them. This type of measurement is known as standard scalar measurement, since each individual instance of it results in a numerical value that refers to a point on some particular standard measurement [conceptual unit] scale, such as: inches, centimeters, degrees Kelvin, degrees Celsius, pints, liters, bushels, grams, ounces, light years, volts, ohms, etc. A standard scalar measurement is a point on a standard measurement scale.

1. **Absolute scale** - When measurement involves simply counting out the number of a set of items or events according to the series of cardinal numbers (i.e., one, two, three, four, etc.), then the scale of measurement is otherwise known as an absolute scale. An absolute scale is a system of measurement that begins at a minimum, or zero point, and progresses in only one direction. An absolute scale begins at a natural minimum, leaving only one direction in which to progress. When counting with cardinal numbers (0,1,2,3,4,5,...), the cardinal number set represents the scale. An absolute scale can only be applied to measurements in which a true minimum is known to exist.
    A. Absolute scales are typically used in science and anywhere precise values are needed in comparison to a natural, unchanging zero point.
    B. Measurements of length, area and volume are inherently absolute, although measurements of distance are often based on an arbitrary starting point.
2. **Relative (or arbitrary) scale** - All other commonly recognized measurement scales are relative, in the sense that they are designed to measure, not the absolute number of items or events, but rather the ‘magnitude’ of some particular attribute (e.g., length, width, weight, temperature, velocity, electrical potential, etc.) relative to the units [of some particular scale that has been designed, or has evolved, for taking the measure of that attribute].

In concern to several important measurable physical quantities:

1. Weight can be absolute, such as atomic weight, but more often they are measurements of the relationship between two masses, while measurements of speed are relative to an arbitrary reference frame.
2. Temperature has a known minimum, absolute zero (where all vibrational motion of atoms ceases), and therefore, can be measured either in absolute terms (kelvins or degrees Rankine), or relative to a reference temperature such as the freezing point of water at a specified pressure (Celsius and Reaumur) or the lowest temperature attainable in 1724 (Fahrenheit).
3. Pressure is a force that can be measured absolutely, because the natural minimum of pressure is total vacuum. Pressure is frequently measured with reference to atmospheric pressure rather than on any absolute scale, relative to complete and perfect vacuum. And, with measurements of things like blood pressure or tire pressure, a measurement relative to air pressure is a better indication of "burst pressure" (damage threshold) than an absolute scale.

3.16 Conceptual composition of 'quantity' by attribute

The concept of a [measurable] 'quantity' is sub-composed of the following attributes (every quantity may be categorized according to these four sub-conceptions):

1. **Measure[able quantity]** (it has a physical referent) – The physical variable. For example, molecular density. If something has a quantity, then it is measurable.
2. **Dimension** (it has a dimension or is dimensionless) - A dimension is a measure of a physical variable (without numerical value[s]). It could be said that a dimension refers to the extent of all possible units of a given type. For example, length is a dimension. The terms measurable quantity and dimension are often synonymous.
3. **Unit** (it has a unit name-quantity) - a unit is a logical process ("way") to assign a number (or, "measurement") to that dimension. A unit is a name (and description) of the value being measured for. For example, the meter is a unit of length.
4. **Value** (it has a numerical value) – the number, logically sequenced to represent the 'magnitude' or 'amount' of an instance of the physical variable (a pattern). For example, when there are three interval patterns of the category meter, then the value is 3.

NOTE: *There are 7 quantities on which all international[ly standardized, ISO] quantities are based. Yet, a system of three base units, consisting of units of mass, length, and time, is sufficient to express the units of all other mechanical quantities.*

3.17 Measureable quantities

NOTE: *The concept 'measure' conveys the idea that measurement involves a chain of coherently connected relationships, which start the sequencing of their relationships at a source, and with which ("against") all new information of a like kind is compared (for patterns). The source of the chain of conceptual relationships is called its: base, fundamental, or axiomatic conception.*

There are two (or at least one, which becomes two) axiomatic/base conceptions of a quantity in science and measurement:

1. **Base/fundamental quantities** are those quantities that are common to any object or event. For example: length, mass, charge, time. A base quantity is a conventionally chosen quantity. No base quantity can be expressed as a product of powers of the other base quantities. Hence, it is said that base quantities are mutually independent (axiomatic).
    - **System of base quantities** – No member of the subset can be expressed in terms of the others. In terms of expression, every other quantity can be conveniently expressed in terms of base quantities. Normally, the symbol of a quantity is written in italics, and that of its dimension in capital letters.
2. **Derived quantities** are quantities formed by combining two or more base quantities (using multiplication or division, algebra). For example, area (length x width), volume (length x width x height, a3), speed (distance/time). A quantity in a system of quantities, which is defined in terms of its base quantities.

Or, the conception of physical quantity could be viewed as follows:

1. Base axiom quantities.
2. Derived mathematical operational quantities.

A complete system of quantities includes both base and derived quantities.

3.17.1 The dimensional attribute of quantity

Dimension refers to the name of the quantity being measured. The "dimensions" of a quantity refer to the basic/fundamental composition ("nature") of the quantity (i.e. how the quantity is related to the base/fundamental quantities of existence: length, mass, time, charge, etc). Every measurement consists of an empirical comparison of dimensions.

In concern to dimensionality, the measurement of a physical quantity may be classified as one of the following:

1. **Dimensionless with units** – has units, but no dimensions.
   - Quantities having units, but no dimensions include, but are not limited to: plane angle, angular displacement, solid angle. These physical quantities possess units, but they do not possess dimensional formulas.
2. **Dimensionless without units** – has no units, and no dimensions.
   - Physical quantities having no units, cannot possess dimensions: trigonometric ratios, logarithmic functions, exponential functions, coefficient of friction, strain, Poisson's ratio, specific gravity, refractive index, relative permittivity, relative permeability. All these quantities neither possess units nor dimensional formulas.
3. **Dimensioned with units** – has units and dimensions.
   - Quantities having both units and dimensions include, but are not limited to: area, volume, density, speed, velocity, acceleration, force, energy, etc.

A dimensionless quantity is a quantity to which no physical dimension is applicable. It is also known as a "bare" number, "pure" number, or a quantity of dimension one. A "pure" number is a number with no unit attached. For example, 2 is a dimensionless quantity, 2 apples, is not (as in, the dimension is "fruit'). Other dimensionless quantities include, for example: 1, i, π, e, and φ, 1/φ, and 1/φ-3 or φ3.

All "pure" numbers are dimensionless and unitless quantities, for example 1, i, π, e, and φ. A "pure" number is a kind of a number that has a dimensionless quantity, and does not have a physical unit. Note here that the use of the word "pure" as a qualification of number, is not useful, because the numeric portion of a dimensioned number is also "pure" in the sense that it is a value.

Presently, given what is known, there are between five and seven primary dimensions (or dimensional [physical] quantities) to material reality. Primary (a.k.a., basic or fundamental) dimensions are defined as independent or fundamental dimensions, from which other dimensions can be obtained. In other words, these dimension are axiomatic to our conception of the dimension of something real.

Hence, remember that when working in mass the base standard to which every other standard and mass measurement is being compared, is the *kilo*+gram (kilogram), and not the *null*+gram (gram). The base unit is a [prefix] multiple of the one unit, kilo-.

3.17.2 Unit attribute

Whereas a quantity is a measurable property for a phenomenon, body or substance, a measurement unit is chosen by convention as the reference to which measurements of that property refer.

The presence of the 'unit' signifies the type of relationship that exists between the number part of a measure (the 'value') and the dimension part of a measure. "Units" refer to specific ways of reporting (or denoting) a quantity.

A unit is the label of a scalar quantity, defined and adopted by cooperation/convention, with which any other quantity of the same dimensional kind can be compared to express the ratio for the two quantities as a number.

> **NOTE:** *The ratio of two quantities of the same dimensional kind is a pure[ly dimensionless] number.*

The measurement unit allows for the [numerical] value [of the quantity] indicated for an object or event, to be compared with the value indicated for the measurements' reference dimension (e.g., mass, length, etc.). The reference dimension is the reference used to calibrate the measurement system (i.e., the relative source of all standard comparisons).

The measuring unit is the relational signifier assigned to the numerical measure, to identify it out of all potential possible [unitized] representations.

Corresponding to a system of quantities, where there are base and derived quantities, there is also a system of corresponding units, where there are base and derived units. A system of [measurement] units is a set of measurement units corresponding to every quantity in the system of quantities. The set of system units consists of:

1. **Base units** (a.k.a. system of base units)
2. **Derived units** (a.k.a., system of derived units)
3. **Dimensionless quantities** (or, quantities of dimension 1)

For every base 'quantity' (as a concept), there exists a base 'unit'. Base units can be used to build and/or express newly "derived" units (Read: derived units). The principal set of units (also sometimes viewed as a subset of units) from which all other units are expressed, is called a [system of] base units.

> **NOTE:** *The magnitude of any given quantity can be expressed by (i.e., associated with) a number equal to the ratio of the quantity to its unit.*

The following are important principled clarification on the conception of base units:

1. There is only 1 base [standardized] unit for a quantity. In every system of units there is only one base unit for each base quantity. For example, in the SI, the metre is the base unit of [the dimensional quantity] 'length'. The centimeter and the kilometer are also units of length, but they are not base units in the SI.
   - **A [unit] conversion factor** – is a ratio of two measurement units of quantities of the same kind. For example, km/m=1,000 and thus, 1km=1,000m; here, 1/1000 is the conversion factor (ratio).
   - In the SI, there are seven base [quantity] dimensions. Mass is the only base/fundamental dimensional quantity whose base [standardized] unit is a conversion factor, the kilogram (Read: kilo as a 1000 conversion factor combined (+) with 'gram' as unit mass). The other base dimensions are standardized to a non-conversion factor base. In other words, notice how length is meter, time is second, amount of substance is mole, but mass is kilogram. Mass is not standardized to a zero-conversion factor.
2. A base unit may serve for a derived quantity of the same dimension. For example, when rainfall is defined as volume per unit area, and the meter is used as a coherent derived unit (in the SI).
3. For any number of entities, the number one, symbol "1", can be regarded as a base unit in any system of units.

   NOTE: *A system of three base units (mass, length, and time) is sufficient to express the units of all other mechanical quantities. However, a system of four base quantities is required to express each and every other quantity.*

A particular quantity "can" be reported in many different kinds of units, but it will always have the same dimensions. It is best to have a unified measurement model where a particular quantity can only be reported with one particular kind of unit (that may be orderly scaled itself), forming a 1 to 1 matching (pairing) between units and dimensions. For example, in early 21st century society, force (which can be expressed F=ma) has dimensions of mass x length/time2. Here, force can be expressed in different units, which leads to confusion and is a sign of a lack of social cooperation and conceptual integration: Newtons, ergs, pounds-cm per square hour, pressure, force, and torque.

3.17.3 Quantities of the same kind

Quantities of the same kind will have the same unit, but two quantity values having the same unit do not have to be of the same kind. For example, the unit of 'mass density' and of 'mass concentration' is kgm-3, but these are not quantities of the same kind. Similarly, the measurement unit of both frequency and activity of radio nuclides is s-1, but they are not quantities of the same kind. The unit in each [conceptually modeled] case is given a unique unit label, namely frequency is hertz (Hz) and the activity of radio nuclides is Becquerel (Bq).

3.17.4 Unit prefixes (unit multiple prefixes)

Because the continuum of each physical unit is so large, notation via multiple(s) becomes necessary for human cognition. A multiple of a unit is indicated by a prefix. The prefixes designating the multiples and submultiples of physical units (e.g., length, frequency, power) are: deca-, hector-, kilo-, mega-, ...

3.17.5 Value attribute (number)

Here, the 'value' is the quantity's numerical association, and the 'quantity value' is the value (number) and unit (reference) together. The quantity's value is the measure's number, and the reference is the measurement unit. To be more specific, a 'quantity value' may be expressed as either:

1. A number and a measurement unit (the unit one is generally not indicated for a quantity of dimension one).
2. A number and a reference to a measurement procedure.
3. A number and a reference material.

A 'quantity' value maintains the following characteristics:

1. The number can be real or complex.
2. A quantity value can be represented in more than one way.

The size (magnitude) of a quantity is expressed as a number accompanied by a measurement unit, and if appropriate, by additional reference to a measurement procedure or a reference material. Here, the term 'quantity value' refers to a number "multiplied by" its tagged unit (or reference), forming a mathematical 'term' upon which calculation operations are possible.

Clarification: The term "true" value is sometimes used in common parlance. The concept of a 'true value' has been redefined in metrology. It used to mean: The true value (of a quantity) is the value which characterizes a quantity "perfectly" defined, in the conditions which exist when that quantity is considered. The concept of a 'true quantity value' now means: a quantity value consistent with the definition of a quantity.

3.18 Conceptual systems model of measurement

Measurement may be represented as a logical [conceptual] information system with inputs, operational processes, and outputs.

1. Inputs [static]: logic, definitions, algorithms, reference standards, units. Answers the question, What is required for measurement?
2. Inputs [dynamic]: quality (property, attribute, characteristic), quantity, measurand. Answers the question, What is being measured?
3. Processes: calibration, measurement. Answers the question, How does measurement occur?
4. Outputs: quantity value (a.k.a., quantity or value), indication, result, term. Answers the question, What is the output of measurement?

3.19 Input (static): Measurement standards

Any measurement requires a measurement standard (Etalon), which is the embodiment of the definition of a given quantity, with state quantity value and associated measurement uncertainty, used as reference.

There is a hierarchical mapping to the concept of a standard [of comparative] reference. The primary (base) standard (in the hierarchy) is checked (i.e., calibrated) to be the same value as new standard at one level lower in the mapping chain. This new, second (and not primary) standard can be used for direct comparison, or a new lower level (third) standard could be checked (i.e., calibrated), and the process continues. All standards are pre-aligned or pre-calibrated, except for the primary, which acts as the standard for all calibrations.

Working/operational level standards are lower in the hierarchy, and that which is closest to a pure conception of the object/event being measured is higher in the hierarchy. Herein, working-level standards are, in turn, calibrated against higher-level standards, which reference (trace, map, or have been demonstrated to align) back to a primary [procedural unit] system.

> **NOTE:** *A proper chain of traceability must include a statement of uncertainty at every step.*

The [principal and principle] standard is the physical representation of the units defined in the system of units.

In measurement, standards define the units and scales in use, which allow for comparison of measurements made in different times and places.

Simplistically, a standard is a process or system that has been agreed upon (formally or informally). Standards emerge when cooperation is present, or when safety is required. Therein exists the necessity for a repeatable process for comparing within a single category some known quantity and/or value with some to be known quantity and/or value.

Standards define that which is required to effectively and cooperatively perform some function (or, "do some thing"). Take note, however, that there are multiple kinds of standards, and their meanings may vary slightly: measurement standards; standards of practice; protocol standards; etc. In general, however, the term means a well-defined and agreed upon set of processes.

Here, measurement standards are "devices" (tools, processes, and objects) that represent the standard system's unit (e.g., SI) in a measurement.

A [measurement] standard is a fundamental reference for determining the value of new information moving into a measurement system. The four common definitions of a measurement standard are:

1. Measurement standards are those devices, artifacts, procedures, instruments, systems, protocols, or processes that are used to define (or to realize) measurement units, and on which all lower echelon (less accurate) measurements depend.
2. A measurement standard may also be said to store, embody, or otherwise provide a physical quantity that serves as the basis for the measurement of a new quantity of some thing [that the measurement standard accounts for].
3. A standard is the physical embodiment of a measurement unit, by which its assigned value is defined, and to which it can be compared for calibration purposes.
4. A standard is a unit of known quantity or dimension to which other measurement units can be compared.

Multiple measurements of a similar category or thing require a 'standard' to which the measuring instrument and/or observer will refer when determining the measure.

> **NOTE:** *Any quantity used as a standard of reference is a unit of measure.*

There are two primary measurement standard categories:

1. **The base/fundamental physical [standard of] reference** – a fundamental physical constant is used as the reference. The real physical world is the metric.
2. **The derived [standard of] reference** – something has been designated by a conscious observer or decisioning system to act as a reference standard. Some object (or digital process) is used as the reference standard. The fundamental physical referent is the reference.

When the word 'magnitude' is used, then there are two measurement system magnitudes:

1. **Fundamental magnitudes (fundamental measurement procedures)** – magnitudes determined from measurement procedures that satisfy the conditions of additivity and do not involve the measurement of any other magnitude. Here, ordering and concatenation operations are active.
2. **Derived magnitudes (derived measurement procedures)** – magnitudes that can only be determined through their relations to other, fundamentally measurable magnitudes. Note here that additivity is not necessary for a measurement.

When the word 'quantity' is used, then there are two measurement system quantities:

1. **Fundamental quantity (a.k.a., base quantity, basic quantity, or metric quantity)** – defined by specifying a[n operational] measurement process.
   - Fundamental base[-ic] system of dimensional units [of measure-ment]
2. **Derived quantity** - Defined by algebraic[ally expressed] combination of base units.
   - Derived/defined units of relational dimensions [of measure-ment]

3.20 Input (dynamic): The measurand (the measured variable)

The measurand is a[n object] name and a description of a (particular) quantity intended to be measured) as the first step in the process of measurement. This step or phase is known as the "problem of definition". The measurand is a description of the specific quantity intended to be measured. The specification of the measurand should be sufficiently detailed to avoid any ambiguity. The measurand is not just another name for analyte. Analyte is the component represented in the name of a measurable quantity, whereas measurand refers to a specific quantity to which quantity values are expected to be attributed by means of a measurement.

The measurand is also sometimes called the 'state name variable'. Technically, that which is being measured is the state of a system, which have been given a [logical] name, and represents a variable (static or dynamic) in the operation of the system. Wherein, more information is being gathered about the system by inquiring into a bounded sub-set of the system.

Hence, a measurand is a particular quantity subject to measurement. Measurand is the [label given in the context of measurement to the] quantity intended to be measured. The measurand is the measured variable.

The problem of clearly defining the measurand ("that which is measured") is called the problem of definition, and it has two parts, one simple and one rather subtle and complex.

The first and simplest part of the phase of definition, relates to the identification (pattern recognition, categorization) of the quantity measured. In principle, all that is required is to provide sufficient information to allow the measurement to be repeated, or otherwise predicted similarly. Herein, influences may also have to be specified.

The second and most difficult part of the problem of definition relates to the technical definition of the attribute that is being measured. In the case of 'temperature', what is the meaning of 'temperature'? The temperature of a system is strictly defined only in conditions of thermal equilibrium, that is no net flow of heat between any of the components of the system.

Measurements with definition problems (i.e., incoherencies) are often the source of great argument. The telltale sign of a definition problem is a measurement where the result seems to vary with the measurement technique.

The standard definition of the term 'measurand' was once different. In fact, the former definition of measurand was, "The quantity subject to measurement." It could be said that a quantity is experimentally 'subject' to the interaction with a measuring instrument (in part, because the measurement instrument is designed and applied with an intent).

3.21 Measurement processes

In addressing measurement problems, it is necessary to have a conceptual model of the measurement process.

3.21.1 Calibration

Calibration is comparison of a measurement device against a standard and adjustment if necessary. Calibration is the process of checked some quantity against a known standard, and adjusting the quantity [to match the known standard] if necessary. Calibration is a comparison of an item to a standard that is closer to the primary (or the primary) standard (e.g., SI), also known as a higher-level standard. Adjustment of the lower-level standardized device is part of the operation required (as in, if the lower-level device does not align correct). Such comparison requires traceability of the calibration. Traceability is defined as an unbroken chain of comparisons to National or international standards (e.g. standards maintained by NIST), AND stated uncertainties at each step. The traceability of course needs to ultimately go to SI.

Calibration is a set of operations that establish, under specified conditions, the relationship between values of quantities indicated by a measuring instrument or measuring system, or values represented by a material measure, or a reference material, and the corresponding values realized by the standard.

Calibration can be further defined:

1. **Calibration** - operation that, under specified

conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication.
2. **Calibration** is the procedure of establishing a relation (adjusting or checking) a scale so that the readings of the (lower-level standard) instrument conforms to an accepted standard (higher-level standard).
3. **Calibration hierarchy** - sequence of calibrations from a reference to the final measuring system, where the outcome of each calibration depends on the outcome of the previous calibration.

The results of a calibration include:

1. The result of a calibration permits either the assignment of values or measurands to the indications or the determination of corrections with respect to indications.
2. A calibration may also determine other metrological properties such as the effect of influence quantities.
3. The result of a calibration may be recorded in a document sometimes called a calibration certificate or a calibration report. (International Vocabulary of Basic and General Terms in Metrology)

In a calibration experiment, the analyst typically prepares a set of calibration solutions (also known as, calibrators, standard solutions, or working standards; i.e., a set of measurement standards). When measured, each of them gives rise to an indication (signal, response). The relation y = f(x) between the indication and the corresponding quantity value is called a calibration curve. The uncertainty of the calibration will include contributions from the uncertainty of the measurement standards, variation in indications, and limitations in the mathematical model when establishing the relation y = f(x).

Calibration and verification are carried out on measurement equipment to determine (and adjust if necessary) their accuracy.

In application, calibration is the process of mechanically or electronically setting the parameters for a measuring instrument.

In concern to calibration, there are two types of measurement instruments (devices, equipment, tools, etc.):

1. **Adjustable** – The instrument's operation may be changed (adjusted) to result in different measurement readings (indications) for the same quantity under the same conditions. Adjustable instruments are calibratable after their creation. Calibration may occur on adjustable measuring instruments after their creation (e.g., micrometers, scales, verniers, etc.). Because the accuracy of adjustable instruments will drift naturally over time, these instruments must be periodically calibrated against a higher-standard.
2. **Non-adjustable** – The Instrument's operation cannot be changed (i.e., it is non-adjustable). Non-adjustable instruments are calibrated once during their creation, which cannot be changed without re-creating the instrument.

Calibration process ("pipeline") involves:

1. Identification
   - Purity of substance
2. Verification
   - Secondary reference materials (RMs) and controls
3. Recognition

A measurement standard is the prerequisite of any calibration, which is the operation that establishes a relation between the quantity value provided by a measurement standard and corresponding device outputs (Read: indications), with associated uncertainties. A calibration may be expressed by a calibration diagram, calibration curve, or calibration table. It can be an additive or multiplicative correction.

> **Note:** *Calibration should not be confounded with the adjustment of a measuring system, sometimes called autocalibration, which is the set of operations (zero, offset, and span or gain adjustment) performed on a measuring system so that it provides prescribed outputs (indications) corresponding to given values of a quantity to be measured. Therefore, in practice, the best performance will be obtained by a first calibration to determine the approximate magnitude of the adjustment needed, then the adjustment, then a final recalibration.*

Calibration is an operation. Under specific conditions, it has a first step that establishes relations between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties. And, in a second step, use this information to establish a relation for obtaining a measurement result from an indication.

A calibration measurement is traceable to NIST when the parameter being measured is clearly defined:

1. **An unbroken chain:** Comparisons from the measurement result reported by a laboratory all the way back to a nationally recognized primary standard (NIST).
2. **Documentation** – Every link in the chain must be

performed according to documented procedures, and the results of these procedures must be documented. This documents the measurement system.

3. **Competence** – Laboratories performing steps in the chain must have demonstrated competence as demonstrated by accreditation to ISO 17025.
4. **Measurement assurance** – The laboratory must systematically establish the status of reference materials and working standards at all times pertinent to a given result.
5. **Measurement uncertainty** – The measurement uncertainty must be determined for each link in the traceability chain, and the measurement uncertainty must be reported for the final measured result. Uncertainty reporting is mandated for ISO accredited calibrations.

3.21.2 Measurement equipment calibration

When uncertainty is relevant (in a decision), the uncertainty reported by a measurement device system (or calibration certificate) is necessary to calculate the uncertainty of measurement.

The calibration system metadata (calibration certificate) must contain specific information to fulfill the purpose of supporting traceable measurement:

1. Identity of device
2. Location of device storage
3. Location of device usage
4. Identity of device user
5. Measurement data

The following may affect performance of measurement equipment and create uncertainty with their reliability:

1. Drift
2. Environmental factors
3. Component age
4. Shock
5. Misuse

For Community Habitat Service System operations, the equipment must be properly:

1. Monitored
2. Maintained
3. Used
4. Stored
5. Transported

3.21.3 Conditions for calibration

Calibration requires traceability. Traceability requires:

1. An unbroken chain of comparison to National or international standards; and
2. Stated uncertainties at each step.

NOTE: *The only way to "prove" that measurements are right (i.e., there is a "right measurement"), is to prove that their uncertainty is low enough to allow the desired conclusions to be drawn from the results, such as whether or not a workpiece meets its specification.*

3.22 Measurement output

Two types of measurement outputs

- **Instrument indications (or "readings")** - these are properties of the measuring instrument in its final state after the measurement process is complete.
- **Measurement outcomes (or "results")** - these are knowledge claims about the values of one or more quantities attributed to the object being measured, and are typically accompanied by a specification of the measurement unit and scale and an estimate of measurement uncertainty.

3.22.1 Uncertainty

Uncertainty is a parameter associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

3.22.2 Properties of the output (result) of measurement

Accurate measurement requires the following principle enablers (and accompanying tools):

1. **Traceability** – enables comparison over time and place.
2. **Uncertainty** – enables meaningful comparison of results.
3. **Confidence** – enables meaningful interpretation of results.
   - Results are only useful when the same pattern ("thing") is compared.
   - Where uncertainty is assessed qualitatively, it is characterized by providing a relative sense of the amount and quality of evidence (that is, information from theory, observations or models indicating whether a statement or proposition is true or valid) and the degree of the understanding.

3.22.3 Traceability

Measurement includes an experimental (primary method) and representational component (additional standards), the latter implying the requirement for metrological traceability.

Metrological traceability is the property of a measurement result, whereby the result can be related to a reference [unit] through a documented unbroken chain of calibrations (comparisons), all having stated uncertainties. Measurability requires [metrological] traceability. Only measurement results are traceable.

There are five sub-conceptions to the definition of [metrological] traceability:

1. An unbroken chain (or relationships)
2. An uncertainty of measurement
3. Documentation
4. Reference to the formal standard (SI units)
5. Calibration intervals

3.22.4 Claiming traceability

The provider of the result of a measurement is responsible for supporting its claim of the traceability of that result or value.

To support a claim, the provider of a measurement result must document the measurement process or system used to establish the claim and provide a description of the chain of calibrations that were used to establish a connection to a particular specified reference. There are several common elements to all valid statements or claims of traceability:

1. A clearly defined particular quantity that has been measured.
2. A complete description of the measurement system or working standard used to perform the measurement.
3. A stated measurement result, which includes a documented uncertainty.
4. A complete specification of the reference at the time the measurement system or working standard was compared to it.
5. An 'internal measurement assurance' program for establishing the status of the measurement system or working standard at all times pertinent to the claim of traceability.
6. An 'internal measurement assurance' program for establishing the status of the specified reference at the time that the measurement system or working standard was compared to it.

The user of the result of a measurement is responsible for assessing the validity of a claim of traceability.

3.22.5 Measurement as a feedback calculation sub-systems

In the context of decisioning, there are [at least] two feedback calculation sub-systems:

1. Measurement systems are used to assess existing entities by numerically characterizing one or more of its attributes.
2. Prediction systems are used to predict some attribute of a future entity, involving a mathematical model with associated prediction procedures:
3. Deterministic prediction system – the same output will be generated for a given input. The output of deterministic models is fully determined by the parameter values and the initial conditions.
4. Stochastic prediction system – the output for a given input will vary probabilistically. Stochastic models involve some inherent probability ("randomness"), wherein the same set of parameter values and initial conditions will lead to more than one output.

Measures and predictions may (sometimes, must) be validated. Validation routines determine if:

1. A measure is valid if it accurately characterizes the attribute it claims to measure.
2. A prediction system is valid if it makes accurate predictions.
3. Validation is defined in ISO 9000 section 3.8.5 as "confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled".
4. The process of ensuring that the measure is a proper numerical characterization of the claimed attribute by showing that the representation condition is satisfied.
5. The accuracy of prediction systems are validated through the process of establishing empirical means (i.e., by comparing model performance with known data in the given environment.

3.23 Measurement operations for the supra-information system

If supra-system information processing occurs, then the value and its accompanying variable become data (polynomial terms) in mathematical operations.

The result of statistical operations performed on numerical measurement values is the production of additional numerical values.

3.23.1 Utilization and/or calculation

The output of a measurement may be useful in itself, or it may have statistical calculations carried out on it to increase the amount of information available.

The two most important and common statistical calculations are:

1. **Averaging** to determine the arithmetic mean.
2. **The standard deviation** for a set of numbers. The standard deviation of a set of measurements is an indication of how much the measurements vary from their average value. The standard deviation is the 'root mean square' of the deviations.

The following procedure will provide the standard deviation (root mean square) of a set of numbers:

1. Square all the deviations from the mean.
2. Add them together.
3. Divide by the number of measurements.
4. Determine the square root.

3.23.2 Validation and Verification

Validation is defined in ISO 9000 section 3.8.5 as:

- Confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled.

Validation ensures measuring processes are aligned. Validation is the process of determining whether functional and/or performance requirements are met.

Validation has a secondary meaning; it also means ensuring that whatever is verified or calibrated is fit for the intended use. Hence, systems can be validated, measurements can be validated, and tools can be validated for intended uses.

Here, measurement validation leads to either a verified or refuted initial measurement(s):

1. Verification of the previous measurement.
2. Contradiction of the previous measurement -- discover that there is a different measurement available.

Verification means verifying something, such as repeating a measurement process or using another method to check that the results are aligned with what was previously measured.

Note here that there is no need to encode the concept of defense, and hence, there is either verification or contradiction, and not, verification or refutation. Refutation comes from the encoding of defense, and contradiction is simply logical incompatibility.

3.24 The method of measurement

Measurement involves processes, which are also known in some applications as methods. As a method, measurement involves the acquisition of [quantity value] information about the properties of objects (or events); and hence, about a larger and more simplex reality.

There are many measurement methods, because existence can be categorized and inquired into in many ways. For different scientific disciplines, there are different methods. For example, the thermoelectric effect is a measurement method of two meanings: temperature, and the infrared spectroscopy of molecular concentration.

> **TERMINOLOGY:** *A '***reference measurement procedure***' is a process accepted as providing measurement results that fit for their intended use.*

3.25 Measurement as "an approach"

Measurement is a set of operations having the object of determining a value of a measurement result, for a given attribute of an entity, using a measurement approach. A measurement approach is a sequence of operations aimed at determining the value of a measurement result. A measurement approach is either:

1. A measurement method,
2. A measurement function, or
3. An analysis model.

3.26 Methodical measurement categorization

Measurement is/involves a system that processes information to answer a question or otherwise determine a relational value. Herein, there are various ways/modes of categorizing the process by which information about an unknown is acquired.

3.26.1 Measurement categorized by the number of [standard] conversions

There are three modes (methods) of measurement, categorically separated based upon the number of conversions present. A "conversion" refers to a 'signal conversion', whereupon a signal from one source is converted by a system into another signal readable by another system (in a network of signal conversion systems). Here, the signal represents information (that may be useful in our fulfillment) about a surrounding real world.

1. **Primary measurement (direct method, no conversions)** – direct observation and comparison. Does not involve a conversion. For example, compare a length of something along a measuring meter stick, and record the observation.
2. **Secondary measurement (1 conversion indirect method)** – involvement of one conversion. For example, the measurement of a thermometer by someone -- a thermometer changes in relationship to its environment (1st conversion), and then, the observer reads the thermometer. Here, the physical system sends a signal to the thermometer

(1st conversion), and the thermometer sends a secondary signal to the conscious observer.

3. **Tertiary measurement (2 conversion indirect method)** – involvement of two conversions. For example, the measurement of a rotating shaft by someone using an electronic display -- a rotating shaft changes in relationship to its environment (signal output), whereupon an electromagnetic system perceives the rotations (1st conversion). The electromagnetic system then outputs the reading as a (2nd) signal to a digital display (2nd conversion). From there, a conscious observer perceives the digital measurement of the rotating shaft.

3.26.2 Measurement categorized by type of comparison

There are two principal methods of measurement; one which involves humans, and one which does not involve humans (or, at least, human involvement is superfluous to the observation and value determination). When a human's sensory system is involved in the comparison (measurement), then the method is known as direct. When human's sensory system is not involved, it is known as indirect.

1. **Direct (human sensation)** – A unknown quantity is visually compared, directly with another of the same pattern. Human senses are necessary for measurement. Here, results are obtained from direct comparison. Because a human is involved, the results are not always accurate.
2. **Indirect (no direct human involvement)** – An unknown magnitude is measured by an instrument with a referentially standardized procedure. The indirect method consists of a chain of [a] synchronously connected devices, which form a measuring instrumentation system. This system generally consists of a detector element to detect, a transducer to transduce, and a memory database unit to indicate or record the processed signal. This system is as accurate as its design and application, and may or may not involve human effort

3.26.3 Measurement categorized by proximity

There are two types of measurement method categorized by contact proximity between the measurement instrument and the thing being measured.

1. With physical contact – instrument is placed in direct physical (less than 1mm) contact with the object. The sensing element is known as the sensory. Here, contact is generally molecular.
2. Without physical contact – instrument is not placed in direct physical contact (less than 1mm) with the object. The sensing element is known as the sensory. Here, "contact" is generally electromagnetic.

3.26.4 Measurement categorized by method sub-type

There are many possible measurement method sub-types, including but not limited to the following:

1. **Absolute/fundamental** –the measurements of base quantities enter into the definition of the quantity being measured.
2. **Comparative** – comparison of the value of a quantity to be measured with a known value of the same quantity.
3. **Null measurement** – the difference between the measurand value and the known value of the same quantity with which it is compared is brought to zero.
4. **Substitution method** – the quantity to be measured is replaced by a known value of the same quantity, so selected that the effects produced in the indicating device by these two values are the same (a type of direct comparison).
5. **Complementary method** – the value of a quantity to be measured is combined with a known value of the same quantity.
6. **Transposition** – the value of the quantity to be measured is first balanced by an initial known value, and then, balanced by another new known value.
7. **Coincidence** – measurements coincide with certain lines and signals.
8. **Deflection** – the value of the quantity to be measured is directly indicated by the deflection of a pointer on a calibrated scale.

3.27 A measurement system

INSIGHT: *A measurement systems indicates the condition of the environment; whereby informed decisions are taken.*

Measurement systems display and/or record an output quantified according (corresponding) to the variable input quantity. The input to the measurement system is the true value of the variable; the system output is the measured value of the variable. Therein, the measurement process may be viewed as a sub-system that generates and outputs information. A measurement system becomes a collection of procedures, gages and operators that are used to obtain measurements.

The function of a measurement system is to provide accurate information about the relative quantity value of a measurand.

The output/result of the interaction between a measuring system and its measured quantity (measurand) is generally termed an ‘indication’ (expressed with a value, the number part, and a unit, the unifying categorical reference part, forming a mathematical term upon which further statistical calculation can be performed).

In an ideal measurement system, the measured value would be equal to the true value. Here, the accuracy of a measurement system can be defined as the closeness of the measured value to the true value. A perfectly accurate system is a theoretical ideal and the accuracy of a real system is quantified using measurement system error E, where

1. E = measured value − true value
2. E = system output − system input

3.28 System of measurement

A ‘system of measurement’ includes a collection of ‘units of measurement’ and rules (logic) relating them to each other.
A system of measurement is a set of related measures that are used to give a numeric value to something.

A system of related measures that facilitates the quantification of some particular characteristic

A “system of measurement” is also known as a “metric” (which is confusing because the system of measurement used as an international standard is called, “the metric system”). There are many systems of measurement, and the metric system is one of those many.

Systems of measurement can be applied describe physical and conceptual systems. Some systems of measurement describe physical systems. Other systems of measurement describe conceptual systems.

A ‘system of measurement’ involves:

1. A group of ‘units of measurement’.
2. The rules relating them to each other.

Systems of measurement are important for the coherent sensing, communication, and construction of systems in the real world.

3.28.1 The metric system of measurement

The International Metric System is an absolute system. Its basic units are the meter, kilogram, and second. It is called an MKS system.

3.29 Elemental composition of a measurement systems

In some cases, the measurement system is made up of only a single component, which gives an output (signal) according to the magnitude of the variable applied to it. However, in most cases, a measurement system is made up of several ordered elements (sub-systems or blocks) between which a signal passes.

It is possible to identify the following types of element (in order), although in a given system one type of element may be missing or may occur more than once:

1. **Transducer (sensing element)** – The element and/or sub-system in contact with the [phenomenological] process, whose output depends in some way on the variable to be measured. Functionally, a transducer is a device that converts a difficult to measure property into a more easily measured property. The transducer often comes into contact with the measured input. A transducer is sometimes referred to as the sensing element. A sensing element is also known as a transducer (sensor). If there is more than one sensing element in a system, the element in contact with the process is termed the primary sensing element, the others are secondary sensing elements. When digital technology is present, generally, the sensing element transduces the input physical effect (input signal) into another physical output, an electrical output signal(s).
    A. The primary sensing element:
        1. Quantity under measurement makes [first] contact with a primary sensing element.
        2. The condition, state or value of the process variable is sensed, by extracting a small part of energy from the measurand.
        3. This element produces an output which maps to (or otherwise, reflects) the condition, state or value of the measurand.
2. **Signal conditioning element** – The element and/or sub-system that takes the output of the sensing element and converts it into a form[at] more suitable for further processing. In most cases, the output of the sensor or the element quantity to be measured is so “small” in signal magnitude that it is not suitable for the output presentation element. The signal conditioning element converts the signal into a form matching the characteristics of the output device (or more suitable for further processing). Common electrical operations performed on the signal here include, but are not limited to: bridging; amplification; oscillation change; and filtration.
3. **Signal processing/conversion element** – The element and/or sub-system that takes the output of the conditioning element and converts it into a form more suitable for presentation. Here, typical calculations are: computational; integrational; and correctional. This is a digital, not analog, process.
    A. Variable conversion (transducer) element:
        1. Map (“convert”) one physical form into

another form [of signal] without changing the information content (meaning) of the signal.
   2. There may be multiple conversions.
  B. Variable modification element:
   1. Modifies the signal by amplification, filtration, or other means so that a desired output is produced according to some mathematical rule.
  C. Data processing element:
   1. Modifies the data before it is displayed or finally recorded.
   2. Performs mathematical operations: To calculate average, statistical, and logarithmic values. To convert data into desired form. To separate undesired signal from noise. To provide correction on the output signal.
4. **Signal utilization element** – The element and/or sub-system that displays the signal to an observer, records the signal, and/or uses the signal as input into a functional control system.
5. **Data presentation element** – The element and/or sub-system that presents the measured value in a form which can be easily recognized by an observer. Common examples of these are: simple pointer-scale indicator (indicator gauge); chart recorder; alphanumeric display; visual display unit; and virtual simulation.
   A. Provides a record or indication of the output.
   B. Transmitting information (measured quantity) - to another location or device.
   C. Signaling – to give a signal that the pre-defined value has been reached.
   D. Recording – to produce a continuous record of measured quantity.
   E. Indicating – to indicate the specific value on a calibrated scale.
6. **Data transmission element** – The system, sub-system, or element, that transmits the signal from one location to another without changing its information contents.

3.30 Instrumentation

Tooling is the automation of the process of measurement capture and computation, and it is desirable if efficiency and optimization is valued. An instrument is a device that transforms a physical variable of interest (the measurand) into a form that is suitable for recording (the measurement).

When using an instrument to take observations of a variable, it is essential to apply/encode the following:

- Validity, unbiasedness, and reliability

Measurement instruments have three primary functional elements:

1. **The detector/sensor** – detects and responds to measurement.
2. **The transducer** – converts measurand to an easier to measure property.
3. **The signal conditioner** – modifies signal.
4. **The readout** – displays result.

3.30.1 Instrumentation reading quality

The following terms are used to describe the quality of an instruments reading:

1. **Range** – The region between the limits within which a quantity is measured, received or transmitted expressed by starting the lower and upper range values.
2. **Span** – The algebraic difference between the upper and lower range values.
3. **Measured variable (a.k.a., measurand)** – a quantity, property or condition that is measured.
4. **Accuracy** – indicates the deviation of the reading from a known value. Accuracy is typically expressed as:
   A. Percentage of full scale reading (upper range value).
   B. Percentage of span.
   C. Percentage of actual reading.
5. **Uncertainty** – Uncertainty of measurement is the doubt that exists about the result of any measurement. Uncertainty is important to make good quality measurements and to understand the results. It is also important in calibration.

3.30.1 Instrumentation systems

An instrument is a device for determining the value or magnitude of a quantity or variable. An instrumentation system is an assembly of various instruments and/or components interconnected to measure, analyze and control various physical quantities (variables). The purpose of an instrumentation (measurement) system is to present an observer with a numerical value corresponding to the variable being measured.

Applications of measurement systems include, but are not limited to:

1. **Monitoring of processes and operations** – measurement systems display and/or record data.
2. **Control of processes and operations** – control systems use measurement data to adjust functioning.
3. **Experimental analysis for science** – science involves the usage of measurement data to evolve/

advance.

When action is taken based on measurement, then the measurement serves a [system] control function. Note that a control system that automatically controls its own functioning based upon its own measurements, is known as, an automatic control system.

Here, there are two principal categories of operation (or processing):

1. **Closed-loop systems (a.k.a., a feedback control system)** – A control system that uses the concept of an open loop system as its forward path, but has one or more feedback loops (hence its name) or paths between its output and its input. The control system measures the value of the parameter being controlled at the output of the system, and compares it to a desired signal, then adjusts its functioning if required. This is also known as an automatic feedback control system, or cybernetic/ cybernated system.
2. **Open-loop systems** - To control the variable, it is first necessary to measure it. Here, an environmental signal enters a controller. The controller is required to:
    - Compare the output variable with the desired value of the controlled variable, and
    - React by sending a message to the control element to take corrective action.

To solve engineering problems two general methods are available: theoretical and experimental. Many problems require the application of both methods. Types of experimental-analysis problems include, but are not limited to:

1. Study of phenomena with intention of developing a theory.
2. Testing the validity of theoretical predictions.
3. Formulation of generalized empirical relationships.
4. Determination of material, component, and system composition.
5. Determination of material, component, and system parameters, variables, and performance indices/ metrics.
6. Solutions to mathematical equations by means of analogies.

In concern to scale, there are two primary types of instrumentation (signal measurement) scaling system:

1. **Absolute systems** – generate and/or measure an absolute signal (e.g., the position).
2. **Incremental system** – counts the number of steps between positions.
3. **Pure** – the number of steps between the start of the system and now is provided.
4. **Referential** – there is a reference, where a reference position is aligned with upon increment of the scale.

The clock is an absolute measurement system, it allows consciousness to determine a point in time. A stop watch is an incremental system, it allows consciousness to determine how many seconds (increments of 'time') have occurred ("gone by") since the start of the measurement.

3.30.2 Sensor / Measuring instrument

A measuring system may consist of only one measuring instrument. A measuring system is one or more (a set of) sensors distributed in space, and an integrated means for data processing information from the sensor with a set of pre-existing, and pre-structured information.

A measuring system must include the following:

1. **Sensor** - A sensor is an element of a measuring system that is directly affected by a phenomenon, body, or substance carrying a quantity to be measured. A sensor is the sensitive element in a measuring system. In practice this term is also applied for designating the device that:
    A. Includes a group of sensitive elements (e.g. an array);
    B. Consists of a number of separate transducers, connected in series (e.g. a primary transducer and amplifier);
    C. Contains a) and b) + additional signal processing units (analogue-to-digital converter, interface, microcontroller, and indicator in any combination).
2. **Intelligent sensor (a.k.a., smart sensor)** - sensors which are able to:
    A. Realize automatic switching of a sub-range of measurements,
    B. Introduce corrections depending on a change of influence quantity,
    C. Carry out automatic self-check of metrological serviceability (a self-validating, adaptive, self-checking, self-diagnosing, self-calibrating, fault-tolerant sensor).

It is possible to describe the measuring system consisting of one sensor (item c) of the list). The sensor may contain a number of sensors according to item b) of the list, each of them containing a group of sensors according to item a) and each sensor of the last group containing a group of sensors according to the VIM-3 definition.

Sensor types include, bit are not limited to:

1. **Multisensor** - A group of sensors perceiving the

same physical quantity (an analog: the tip of the tongue.
2. **Polysensor** - A group of sensors perceiving various physical quantities (an analog: a surface of the tongue as a whole).

Sensors are often transducers in that they are devices that convert input energy of one form into output energy of another form.

Sensors can be categorized into two broad classes depending on how they interact with the environment they are measuring.

1. **Passive sensors** do not add energy as part of the measurement process but may remove energy in their operation.
2. **Active sensors** add energy to the measurement environment as part of the measurement process.

Sensory fusion is a process where two or more sensors are used to observe the environment and their output signals are combined in some manner (typically in a processor) provide a single enhanced measurement. This process frequently allows measurement of phenomena that would otherwise be unobservable.

> **NOTE:** *Biological sensor systems enable determining a value of a "measurand", and moreover, evaluating the distribution of the "measurand" in a multiparameter field and forming a "multiparameter image".*

> **INSIGHT:** *The ultimate purpose of intelligence is to ensure the survival of its carrier.*

3.30.3 Measurement system sensor types

Measurement systems sensors may be divided into three categories of guidance.

1. **Direct [guidance]** – wire guidance, magnetic guidance. These are the most reliable. These systems suffer from the considerable problem of path planning. If the path has to be changed, a certain number of hours are required to install the cable inside the floor and the guidance system must be stopped during installation.
2. **Relative [guidance]** - The relative or dead-reckoning methods, such as encoders, gyroscopes, ultrasound, etc., have the considerable advantage of being totally self-contained inside the system, relatively simple to use and able to guarantee a high data rate. However, since these systems integrate relative increments, errors grow considerably over time.
3. **Absolute [guidance]** - use of external references to achieve an absolute measurement with respect to the environment in which the system exists. These systems are more complicated than the relative ones, work at a slower rate, and lead to the problem of the visibility of the targets needed during the systems' path through an environment. Generally, since these measure the system's position and attitude with respect to absolute references (targets), the error is always bounded and absolute repeatability guaranteed.
4. **Combination** – a combination of the three types.

From the above considerations it is clear why many systems currently make use of both a relative and an absolute system.

3.31 Measuring devices, instruments, and tools

Sensors observe, sense, and otherwise, interact with the environment. Sensors observe stimulus, producing an "observation" of a property(s), such as time, location, and distance. And, this observation allows for interactive change with the environment generating or otherwise relating to a stimulus.

A measuring instrument is a device used for making measurements alone or in conjunction with one or more supplementary devices (as part of measuring system). A measuring instrument (gauge) is frequently a form of transducer. A transducer is a device that provides an output quantity (most often an electric current) having a specific relation with an input quantity (most often a physio-logic signal). The physiologic signal is collected by a sensor defined as an element of a measuring system that is directly affected by a phenomenon, body, or substance carrying a quantity to be measured, or less frequently, by a detector defined as a device or substance that indicates the presence of a phenomenon, body, or substance when a threshold value of an associated quantity is exceeded.

3.32 Properties of measuring devices

Measuring instruments provide a quantity value for the measurand. Therein, the measuring 'interval' or measuring 'range' is the set of values of quantities of the same kind that can be measured by a given instrument with specified instrumental uncertainty under defined conditions. A measuring instrument/system is characterized by [at least] the following properties or quality criteria:

1. **Sensitivity** - the minimum change in input signal to which an instrument can respond. Sensitivity is the ratio of o/p response to a specific range in i/p.
2. **Selectivity** – is a property used with a specified measurement procedure, whereby it provides measured quantity values for one or more

measurands, such that the values of each measurand are independent of other measurands, or other quantities in the phenomenon, body, or substance being investigated.

3. **Resolution** - is the smallest change in a quantity being measured that causes a perceptible change in the corresponding output of the measuring instrument.
4. **Stability** - is the property of a measuring instrument, whereby its metrological properties remain constant in time. An instrumental drift is the continuous or incremental change over time of the indication because of change in metrological properties.
5. **Step response time** - is the duration between the instant when an input quantity value of a measuring instrument is subjected to an abrupt change between 2 specified constant quantity values and the instant when a corresponding device output settles within specified limits around its final steady value.
6. **Maximum permissible measurement error or limits of error** - the boundary value of measurement error, with respect to a known reference quantity value, permitted by specifications for a given measurement, measuring instrument, or measuring system. The term 'tolerance', which is the magnitude of permissible variation of a quantity, should not be used to designate the maximum permissible error. Tolerance includes the true value ± the maximum permissible error.

Strictly speaking, accuracy, trueness, and precision are qualifying measurements, whereas sensitivity, selectivity, resolution, stability, and step response time are qualifying dynamic outputs of devices.

> **NOTE:** *The dynamic output of a measuring instrument is sometimes known as an 'indication'.*

3.33 Measurement performance characteristics (i.e., measurement output parameters)

Measurement generally occurs through instrumentation. That instrument has a set of output/performance characteristics.

3.33.3.1 Static performance characteristics

Static performance characteristics include principally, that desired input to the instrument not change in relation to time. Therein, the following sub-conceptualizations are required:

- Error, accuracy, calibration, hysteresis, dead zone, drift, sensitivity, threshold, resolution, precision, repeatability, reproducibility, linearity, etc.

3.33.3.2 Dynamic performance characteristics

Dynamic performance characteristics include, but are not limited to: speed of response, measuring lag, fidelity, frequency response, dynamic error, overshoot, dead time and dead zone.

Therein,

1. **Readability** indicates the closeness with which the scale of the instrument may be read.
2. **Least count** is the smallest difference between two indications that can be detected on the instrument scale.
3. **Range** represents the highest possible value that can be measured by an instrument, or limits within which the instrument is designed to operate
4. **Linearity** is a measurement system category; wherein, a measurement system is linear if the output is linear proportional to the input.
5. **Repeatability** is the ability of a measuring system to repeat output readings when the same input is applied to it consecutively, under the same conditions and in the same direction. Repeatability is expressed as the maximum difference between output readings.
6. **Reproducibility** is the degree of closeness with which the same value of a variable may be measured at different times.
7. **System response** is the ability of the system to transmit and present all the relevant information contained in the input signal.
8. **Threshold** is the minimum value of a ratio (e.g., i/p or a/b) required to cause a detectable change [from 0(zero)], o/p.
9. **Hysteresis** is the maximum differences in two output (indicated values) at same input (measurand) value within the specified range when input is continuously increased from zero and when input is continuously decreased for maximum value.

3.34 Measurement uncertainty and error

Error is the difference between the measured value and the 'true value' of the thing being measured. Uncertainty is a quantification of the doubt about the measurement result. Here, the accuracy of an instrument is defined as the difference between the true value of the measurand and the measured value indicated by the instrument. Typically, the true value is defined in reference to some absolute or agreed upon standard.

A measure (attribute) is "well-defined" if scale and unit are clearly specified; specification of the unit and scale ensures the measure is unambiguous.

3.34.3.1 Measurement accuracy

Measurement accuracy is the closeness of agreement between a measured quantity value and a true quantity value of the measurand. The concept of accuracy is a quality and is not given a numerical value. A measurement is said to be more accurate when it offers a smaller measurement error. Therefore, a measurement error is qualifying a single measurement.

3.34.3.2 Measurement trueness

Measurement trueness is the closeness of agreement between the average of an infinite number of replicate measured quantity values and the true or a reference quantity value. The concept trueness is a quality and is not given a numerical value. Measurement trueness is inversely related to systematic measurement error but not to random measurement error. Since the mean random error is zero, the bias (average of measured value – reference value) is an estimate of the systematic measurement error. The traditional averaging of (measured value – reference value) is equivalent to the former formula only when there is a unique true (or reference) value. When there are different quantities of the measurand, the bias can be fixed, proportional, or distributed following specific functions. Since a systematic error cannot be normally/randomly distributed, averaging (measured value – reference value) is therefore an approximate representation of the averaged bias (systematic error).

3.34.3.3 Measurement precision

Measurement precision is the closeness of agreement between measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. Measurement precision is related to random measurement error and usually expressed numerically by measures of imprecision, such as standard deviation (σ), variance (σ2), or coefficient of variation (σ/mean) and assuming a mean = zero.

3.34.3.4 Measurement error

Measurement error. We have seen in "Measurement Accuracy" that it is the difference between a unique measured quantity value and a reference quantity value. The measurement error can be systematic (bias) qualifying the untrueness (i.e., "measurement trueness") or random qualifying the imprecision (i.e., "measurement precision"). When the term "measurement error" is used without further information, it combines systematic and random errors and qualifies the inaccuracy.

- % error = |experimental-accepted| x 100 / accepted value

For any particular measurement there will be some error due to systematic (bias) and random (noise) error sources. The combination of systematic and random error can be visualized by considering the analogy of the target.

There are three classifications of measurement error:

1. **Gross errors or mistakes**
   A. **Accidents and mistakes** – can avoid only by staying focused (and in flow).
2. **Systematic error sources (bias)** – these have definite magnitude and direction.
   A. **Instrument error** – the instrument's design, maintenance, selection, calibration, and/or operation-usage cause an error.
   B. **Environment error** – due to conditions external to the measuring instrumentation sub-system. Use the instrument under the conditions (parameters) it was designed and calibrated for. Calibrate for local conditions.
   C. **Observation error** – error due to poor capabilities and/or motivations of operator.
   D. **Operational error** – misuse of the instrument or poor operational technique.
   E. **System interaction error** – an interaction error may occur between a system (to be measured) and the direct point of contact with the instrument's body. The contact could change the condition of the measuring instrumentation system. For example, a ruler pressed against a body (system) resulting in the deformation of the body and a reading with a large error.
3. **Random error sources (noise)**
   A. If systematic errors can be removed from a measurement, some error will remain due to the random error sources that define the precision of the measurement. Random error is sometimes referred to as noise, which is defined as a signal that carries no useful information.
   B. Presently uncategorizable errors.

3.34.3.5 Stability and consistency

All measurements have to exhibit two basic characteristics: stability and consistency. To the degree that they do, they are called, "reliable measures".

1. **Stability** - A stable measure will yield identical measurement results whenever it encounters an identical amount of the theoretical concept.
2. **Reliability** – An operational definition describing a measurement procedure which behaves in a consistent fashion.
3. **Validity** - The amount of measurement validity cannot be determined by any numerical method.

It relies on a self-evident overlap between "verbal world" theoretical definitions and "measurement world" operational definitions. Furthermore, the self-evident overlap must be generally agreed upon by independent observers. It is not enough that an operational definition shows measurement validity to the researcher who constructs it; it must also exhibit the same measurement validity to other researchers and critics.

A. Concurrent validity
B. Discriminant validity
C. Construct validity

3.34.3.6 Measurement uncertainty

Measurement uncertainty is a parameter characterizing the dispersion of the quantity values being attributed to a measurand based on the information used. This concept is broader than precision including uncertainty due to time drift, definitional uncertainty, and other uncertainties. There are 2 types of evaluation of the uncertainty of measurements: type A based on statistical analysis of measured values and type B based on other means such as certified references, authoritative published values, or personal experience.

3.34.3.7 Error triangulation

This process of combining information from multiple sources to arrive at a true or at least more accurate value is called triangulation, a loose analogy to the process in geometry of determining the location of a point in terms of its relationship to two other known points. The key idea behind triangulation is that, although a single measurement of a concept might contain too much error (of either known or unknown types) to be either reliable or valid by itself, by combining information from several types of measurements, at least some of whose characteristics are already known, we can arrive at an acceptable measurement of the unknown quantity. We expect that each measurement contains error, but we hope it does not include the same type of error, so that through multiple types of measurement, we can get a reasonable estimate of the quantity or quality of interest.

3.35 Categorization and classification

Categorization and classification mean essentially the same thing (the process is also sometimes called "division") – the delineation of discreteness of variables. Discreteness of variables is a key part of evaluation and the naming of variables for quantitative research/measurement. This delineation results in an organized structuring of information based upon inter-relationships.

NOTE: *Classification can only occur on similar items (i.e., items expressing a similar pattern).*

In general parlance, however, the words categorization and classification have slightly different definitions:

1. In general parlance, **categorization** is a systematic method of modeling using related conceptual nodes. Every node in the model/map represents a concept, and sub-/supra-nodes represents a more specific/more general concept.
2. In general parlance, **classification** is the representation of knowledge (or awareness) by discrete organization.

Regardless, terms mean the use of concepts (and sub-/supra-concepts) to formalize a difference or similarity (a delineation or division).

INSIGHT: *How someone categorizes and the results of that categorization will affect someone's thinking, and consequently, their motivations and behaviors.*

Therein, a category or class is a collection of similar objects or entities, which are dissimilar from other objects or entities [of a different category or class]. From the perspective of unification, those other categories/classes may exist:

1. At the same level in a leveled/tiered arrangement of categories/classes.
2. At different levels in a leveled/tiered arrangement of categories/classes.

Suffix: -ization and -ication mean "the act of".

1. **Classification** – act[ion] of classifying – identifying similar and different classes.
2. **Categorization** – act[ion] of categorizing – identifying similar and different categories.

The nodes (concepts) in the organized structure are called by different names, including but not limited to: class, type, set, or taxon, and to a lesser degree of usage, unit, topic, and subject.

In other words,

1. Each node in a classification structure is called a class.
2. Each node in a categorization structure is called a category.

The supra- and sub-nodes (supra-/sub-concepts) in a categorization or classification structure (scheme) may be said to exist on different levels or ranks. Rank is the relative level of a group of classified objects/events. A given rank subsumes under it less general categories, that is, more specific descriptions. Above it, each rank is classified within more general categories. Therein,

1. Each level "down" in the structure zooms in on a

smaller and smaller physical/conceptual area.
2. Each level "up" in the structure zooms in on the larger physical/conceptual area.

NOTE: *All measurement data is categorical data. Categorical data is data that can be organized into mutually exclusive categories. Given the presence of categorical information, quantitative and qualitative data may also be considered categorical data.*

3.36 Classification operations

In order to classify/categorize, the following classificatory operations (classification operations) must be performed:

1. Conceptualization operation (defining, identifying terminological and definitional classes/types/categories) – ensuring that concepts (classes or types) have clear, logical, and delineated/bounded definitions. Concepts must be clarified and denominated by a "suitable" (linguistically precise) term or expression. If there is a hierarchy, then concepts "higher up" the hierarchy represent greater unification. Concepts corresponding to individual classes are either formed or clarified by the definition of their boundaries with contiguous concepts.
2. Conceptual analysis operation (divisioning, conceptual elaboration) - the extension of a concept [at a given level of generality] is subdivided into several (two or more) narrower extensions corresponding to as many concepts at lower level of generality; this subdivision is obtained by stating that an aspect of each of the latter concepts is a different partial articulation of the corresponding aspect of the higher concept. Notice that in principle all other aspects of the higher concept are carried into each of the lower concepts. This is a process of conceptual elaboration.
3. Pattern recognition operation (grouping) - the objects or events of a given set are grouped into two or more subsets according to the perceived similarities of their states on one or (more frequently) several properties; subsets may be successively grouped into subsets of wider extension and higher hierarchical level.
4. Assignment operation (assigning, classing, assigning to a class/type) - whereby objects or events are assigned to classes or types which have been previously defined. This is the assignment of objects/events to classes, types or taxa which have been previously defined.

These classification operations produce a classification (categorization) structure that organizes related information by named relationships. That organized information may be viewed as a model/map and used for creating in the world. The input of additional related information becomes organized according to the classification model (which may be known as an ontology, taxonomy, or typology), and its integration may change may change the model (if defining, divisioning, and grouping continue during the assigning process).

3.36.1 Combinatorics

In order to combine parts (or elements) into a whole, a series of processes (process functions) must occur. The three primary functional process to combing parts are:

1. Labeling
2. Selecting
3. Grouping [given all information is known].

3.37 Classification/categorization output: Ontology, taxonomy, or typology, or typology

The result of categorization/classification is an organizational structure known as an ontology, taxonomy, or typology. In common usage, the words typology, taxonomy, and ontology mean essentially the same thing - the categorization/classification of something (or things) and its (their) resulting structurally organized output as an information system (model/map). They are different words used by different disciplines that mean essentially the same thing. In some disciplines, the term ontology is used to imply a broader scope of categorized information about reality and the nature of existence. In other disciples and contexts, the word taxonomy or typology is used.

When ontology and taxonomy are used in the same context, but to mean different things, then the term ontology is likely being used to encompass a number of taxonomies, with each taxonomy organizing a concept/subject/topic in a particular way.

In their disciplined application, the terms have different originations and a slight variance of definition, but still mean the same thing – a structural organization of those objects/entities which have undergone categorization/classification:

1. The word "**ontology**" comes from philosophy, and is a series of categorized characterizations of the nature of being, or reality. Philosophically speaking, it is intended to be a systematic account of existence. Ontologies are concept specifications and relations about reality. The term ontology means the science or study of being and the nature of existence. Etymologically, the word "ontology" comes from modern latin, ontologia (c. 1600), from

onto- + -logy. Onto- means "a being, individual; being, existence.

2. The word "**taxonomy**" means the science of classification. The word was neologized ("coined") by someone studying botany. Etymologically, the word "taxonomy" comes from French taxonomie (1813), coined irregularly from Greek taxis "arrangement" + -nomia "method," from -nomos "managing," from nemein "to manage". The word has come to mean the science of defining groups on the basis of shared characteristics, and giving names to those groups. It is the classification of existence according to characteristics; it is the science of organizing existence into a system of different groups according to the features that they share, and of giving them names. In other words, a taxonomy is a method of partitioning (with purposeful and identified parameters) and giving names.
   A. A taxonomy is a semantic hierarchy in which information entities are related by either the subclassification of relation or the subclass of relation. Note: subclassification is semantically weaker than subclass of relation. A taxonomy is a form of classification scheme. Designed to group related things together. It is a hierarchical thesaurus with terms applied at the final node. It is a tangible hierarchy forming a structure of information related to a root or axiomatic conception. It expresses similar relationships between things. It is an information model. A taxonomy is a knowledge organization system. It is a hierarchy of relationships. It is a hierarchy of related types.
   B. The dimensions of a taxonomy represent empirically observable and measurable characteristics. It is a classification structure based on the empirical/operational.
   C. Each node in a taxonomy may be called a taxon (plural, taxa).
   D. A taxonomy is:
   E. Organized into a hierarchy.
   F. Each tag is unique.
   G. The tags relate to one another logically, and preferentially relate to the existent, which they are classifying.
   H. A structure for organizing incoming information.
3. The word "typology" means the study of types (e.g., types of systems, for example, biological, chemical, linguistical, architectural, etc.) It refers to the science of classifying existence. The study or system of dividing a large group into smaller groups according to similar features, qualities, and characteristics. Typology involves the process of partitioning (based on identified parameters) for the purpose of study. Working with typologies contributes decisively to forming concepts, exploring dimensionality, establishing measurement categories, and grouping cases.
   A. The dimensions of a typology represent concepts rather than empirical classes. It is a classification structure based on the abstract (i.e., the conceptual).
   B. Each node in a typology is called a type.
   C. Typification – act[ion] of typifying – identifying similar and different types.

In computer science, what "exists" is what is represented. Hence, programmatically speaking, ontologies/taxonomies/typologies represent explicit domain conceptual specifications -- an ontology/taxonomy/typology is a domain of interest's formal, explicit specification of a shared conceptualization. Thus, ontologies/taxonomies ensure a shared conceptualizations formal specification.

> **NOTE:** *Tabulation is the logical operation of [numerically] counting the number of cases that fall into each category.*

3.38 Real world category continuity

Real world categories differ in their range and level of continuity. For example, members of the category "atomic element" are highly similar (having atoms, electrons, and protons), and thus, have a relatively small category range and are fairly cohesive (i.e., highly continuous). On the other hand, the category "weapon" is highly variable and contains items such as knives, bombs, guns, etc., which are highly discontinuous.

3.39 Qualitative data type classification

Qualitative data can be nominal and ordinal, but not interval or ratio. Qualitative data cannot be continuous. Discrete, quantitative counts must be ratio. Therein, there are five different possibilities, which forms the following taxanomical concept map:

1. **Categories**: Qualitative, discrete, nominal data (such as colors, names, or labels).
2. **Ranks**: Qualitative, discrete, ordinal data (such as sizes, preferences, or grades).
3. **Counts**: Quantitative, discrete, ratio data that count something.
4. **Relative (or Relative Scale) Measures**: Quantitative, continuous, interval data (such as temperatures).
5. **Absolute (or Absolute Scale) Measures**: Quantitative, continuous, ratio data (such as heights and weights).

3.40 The conceptual components of a measurement system

The measurement of a property [of a bounded existence] may be categorized according to the following sub-concepts (i.e., "criteria"). For a measurement to represent an unambiguous comparison, all four conceptual criteria must have accurate information. The measurement of an attribute (property) of a real world system requires the following conceptual information:

1. Typological/taxonomical positioning (level/scale of measurement): The type or level of measurement is a taxonomy (classification) for the methodological character of a comparison. In other words, the data collected on a variable, and accompanying data, fit into one of several taxonomical scale (or scaling) categories, which determine the methods (and operations) that may be used in its [mathematical-statistical] processing.
2. Numerical quantity (magnitude determination): The magnitude or [numerical] quantity is the numerical value of the characterization, usually obtained with a suitably chosen measuring instrument (a referencing process or tool). A numerical value is a real number that represents a quantity.
3. Unit (referential unit system): A unit assigns a mathematical weighting factor to the magnitude that is derived as a ratio to the property of an artefact used as a standard or a natural physical quantity. Measurement always includes units - without units, a quantity and its corresponding measurement carry no understanding.
4. Uncertainty (determination): An uncertainty represents the "random" and systemic errors of the measurement procedure; it indicates a confidence level in the measurement. All measurements have some degree of uncertainty associated with them, which is usually expressed as a 'standard error of measurement'. Errors are evaluated by methodically repeating measurements and considering the accuracy and precision of the measuring instrument:
    A. Measurement accuracy - How close a measurement comes to the true value (a.k.a., correct value).
    B. Measurement precision - How close a series of measurements are to one another.
    C. Precision refers to how small an uncertainty the measuring instrument and conditions will provide.

3.40.1 Typological/taxonomical positioning(i.e., levels/scales of measurement)

All measurement data and variables fit into one of several possible levels (or scales) of measure[ment]. Together, the levels/scales form a taxonomy/typology known as a 'level of measurement' or 'scale of measure', which classifies the nature of information assigned to a variable.

In specific, the terms, 'level/scale of measure[ment]', refer to the degree of relationship among the values that are assigned to the attributes for a variable upon which measurement data is being collected and will be processed. Associated with each "level" of measurement is a set of permissible [mathematical] transformations [of the data]. (see illustrator for "level of measurement")

In its expression, a 'level/scale of measure[ment]' is a typology (categorically mapped arrangement of concepts) for defining data processed by measurement operations as part of a variable. Data represented as numbers can be grouped/categorized into 4 types (or levels) known as the levels of measurement (or scales of measure). The levels/scales have an order. Each ascending level possesses the characteristics of the preceding level, plus an additional quality.

In common parlance, the terms 'levels of measurement' and 'scales of measure' convey the same (or highly similar) meaning. However, in practice, the terms 'level of measurement' and 'scale of measurement' may have slightly different meanings:

1. **Level of measurement** refers to the particular way/order [in the taxonomy] that a variable is measured, and
2. **Scale of measurement** refers to the particular tool/process for sorting the data that applies based on the level.

Note that sometimes these levels/scales of measurement are referred to as "levels of measurement scales". Regardless of labeling, the concept refers to the classification/categorization of the type of data it is possible to collect from a variable due to the presence of underlying relationships.

'Level of measurement' or 'scale of measure[ment]' is a classification that describes the nature of information within the numbers assigned to variables. It could also be said that [measurement] scales are distinguished by their level of measurement. There are four levels of measurement.

The most commonly used 'level of measurement' typology has three scales of measurement, and one level of basic categorization:

1. Level 1 (not a scale): Basic categorization/sorting – **nominal categorizing**
    - The initiation of categorization; initiating categorization.
2. Level 2 (1st scale): A one dimensional ordering – **ordinal scaling**
    - Associating an ordering dimension to the

categorization.

3. Level 3 (2nd scale): Subdivisioning/delineating the order – **interval scaling**
   - Divisioning the categorical dimension.
4. Level 4 (3rd scale): Affixing/absoluting the order [to a zero point] – **ratio scale**
   - Absoluting the categorization.

> **Note:** *The "list" of available levels/scales [of measurement] is itself an ordinal typology categorization of data complexity.*

There is a [categorical] hierarchy (order) implied in the concept, 'level of measurement'. At lower 'levels' of measurement assumptions tend to be less restrictive and data analyses tend to be less sensitive. At each level up the hierarchy, the current level includes all of the qualities of the one below it and adds something new. In general, it is desirable to have a higher level/scale of measurement (e.g., interval or ratio) rather than a lower one (nominal or ordinal).

Hence, it could be said that measurement data is distinguished by the relationship complexity of the information it carries. All measurements must take one of four (sometime five) forms, also known as "levels of measurement". The four levels of measurement are:

1. **Nominal scale/level (non-metric or categorical)** - lowest level of information. There are only categories (strictly qualitative information). Herein, there is only the assignment of numbers. This is commonly referred to as the "nominal [measurement-level] scale". Note that this scale/level does not represent true measurement.
2. **Ordinal scale/level (non-metric or categorical)** - a higher level of information, ranking scale; it consists of a set of categories that are sequentially ranked-ordered with care to size or magnitude of difference between different variates (quantitative order exists among them). Ranking/ordering of the available numbers by conceptual criteria, but no information is available to derive an understand of how far apart they are separated conceptually and/or numerically. This is commonly referred to as the "ordinal [measurement-level] scale".
3. **Interval (cardinal) scale/level (metric)** - an even higher level of information; interval scale consists of ordered categories with precisely equal intervals between each category. Differences will reflect relative changes in magnitude, but ratios are not meaningful due to the absence of an absolute zero reference point. A scale which represents quantity, and has equal units, but for which zero represents an additional point of measurement is an interval scale. This is commonly referred to as the "interval [measurement-level] scale".
4. **Ratio scale/level (metric)** - the highest level of information. The ratio scale is a specific interval scale that includes an absolute-zero point such that ratios of data do reflect changes in magnitude, precisely. A[n interval] scale that has an absolute zero (no numbers exist below the zero). This is commonly referred to as the "ratio [measurement-level] scale".

Objective measurement scales are otherwise known as numeric measurement scales. The [objective/numeric] scales of measurement are nominal, ordinal, interval, and ratio. These scales represent an order in themselves -- the scales themselves are a scale of increasingly bound understanding. The ratio (or absolute) scale is the most restrictive of all, and the nominal scale is the least bound (and may not even be considered a scale, because it does not scale along any dimensions). The four scale types are ordered in that all later scales have all the properties of earlier scales— plus additional properties.

Each level of measurement and its corresponding scale is able to measure one or more of the four properties of measurement:

1. Identity (nominal)
2. Magnitude (ordinal)
3. Equal intervals (interval)
4. Minimum value of zero (ratio or absolute)

> **HISTORICAL NOTE:** *These levels and scales of measurement were invented by Stanley Smith Stevens, who wrote about them in a 1946 article in Science, titled "On the Theory of Scales of Measurement."*

All levels of measurement give the ability to determine the presence or absence of some thing. A second level of measurement adds the idea of quantity (e.g., "more of ..." or "higher than ..."), or an underlying dimension, to the measure's ability to detect. If the measurement contains only detection and comparative ordering information, then it is called ordinal. At the next higher level, measurement adds the idea of units, such that absolute statements (rather than comparative) can be given about the similarity or difference between measurements. That is, it is possible to state the number of units by which observations are measured to be different. This level of measurement is called interval. Finally, if the measure is interval, but also contains an absolute zero category or scale point, then it is possible to give statements of proportion ("only one half of") or ratios ("twice as much...") about the magnitude of the measurement. This highest level of measurement is called the ratio-level.

The importance of the measurement level is threefold:

1. It determines the selection of test statistics.
2. It affects the amount of information collected about

variables.
3. It affects how questions (inquiries) are formed.

Knowing the level/scale of measurement facilitates decisioning on the interpretation of data from that variable. If a variable is known to be nominal, then it is known that the numerical values are simply short codes for the longer names. Second, knowing the level of measurement facilitates decisioning in concern to what statistical analysis is appropriate on the values that were assigned. The type of scale used in taking measurements directly impinges on the statistical techniques which can legitimately be used in the analysis.

In statistics, different types of data utilize different scales. In other words, not every attribute or variable can be translated to numerical values in the same way.

The nominal scale offers the least statistical information content, and the ratio scale the highest. Nominal and ordinal are non-metric (or categorical) scales; that is, their response values are not directly usable as a numerical value. Interval and ration scales are metric scales that allow for various arithmetic operations.

The level of measurement expresses, how quantifiable a data value actually is (i.e., to what extent mathematical operations can be applied).

Scales [of measurement] with greater complexity, and hence, more categorical data, allow for greater mathematical (logical) processing (analysis and synthesis). In other words, the accurate processing of greater complexity will facilitate greater understanding. A greater number of meaningful operations (data processing) can be done on complex data. This delineation of scales [of measurement] by data complexity is otherwise, and unnecessarily, referred to as levels [of measurement].

The type of measurement scale determines:

- How measurement data is processes.
- Whether statements involving measurement data are meaningful.

TERMINOLOGY: *A continuous scale has units of measurement that are in principle infinitely divisible, so that any particular outcome (continuous variable) could be drawn out to as many points (e.g., decimal places) as practical.*

3.40.2 Scaling

A scale is a conceptual technique to measure some thing. It is an abstract measurement tool for comparing (relating) common attributes of entities. Therein, scaling is the process of ordering a series of items along some type of continuum. A concept, object, or event may be assigned a [measured] number along a scale representing a dimension (or concept). In other words, the data of an observation/sensation is encoded via a rule/principle along a pre-existing comparison continuum known as a [dimensional] scale. A scale is defined as the collection of attributes used to measure a specific [conceptual] variable (e.g., time, temperature, gender, etc.). A scale is a structure for mapping. A particular way of assigning numbers or symbols to measure something is called a scale of measurement (sometimes also called a system of measurement).

Clarification: *In drafting, architecture and engineering, the term scale has two meanings. A scale is a dimension that represents the structure shown in a plan. A scale is also a ruler used in drawing and measuring architectural and engineering plans. Hence, the term scale is sometimes used to refer to a measuring instrument, and sometimes even the standard of measurement.*

Simply, a scale is a rule (principle) used for the assignment of numerals to properties of objects or events. The concept depicting the rule upon which a scale is based is sometimes called the 'dimension' [of the scale]. This equates a scale to a specific method of measurement. Measurement always occurs in a specific way, which means that every measurement process must have a rule of measurement. Every process of measurement must have a scale of measurement.

Note: *A 'measurement scale' is a set of predefined symbols or values in order to represent certain common measures.*

Visually, a scale is a set of points on a line (or, ordered attributes of a concept) used for measuring (associating objects/events with words and/or numbers in a logical manner so that the data can be processed by a mathematical system.

A 'scale' represents the way a variable is measured or quantified. For example, the variable "gender" is commonly measured on a scale defined by the specific attributes "male" and "female". A scale could be considered a "technique" to measure something and integrate it within a numeric, semantic, or graphical system. These are simply ways to categorize different types of variables.

Scaling is the procedure for the assignment of numbers (or symbols) to a property of events/objects in order to impart some of the characteristics of numbers to the properties in question. It describes the procedure of assigning numbers to concepts – a scale is a continuum, consisting of the highest point and the lowest point. Simply, scaling is the assignment of objects to numbers according to a rule. The objects can be linguistic concepts, or numerical concepts.

Scales are generally divided into two broad categories: unidimensional and multidimensional. A scale can have any number of dimensions in it. What's a dimension? Think of a dimension as a number line. If we want to measure a construct, we have to decide whether the concept can be measured well with one number line or whether it may need more.

NOTE: *Concept mapping is a technique for visualizing scales.*

The scale determines what operations among the numbers assigned in a measurement will yield results significant for what is being measured. In other words, it carries the information for an initial interpretation of the numbers arrived at in a measurement. What mathematical transformations measurement can be subjected to depends on the scale in terms of which they were arrived at. Of course, this depends entirely on the available mathematics. There is no advantage in using a scale that allows operations that are no known how to perform.

Scales are important because they define the nature of information about variables.

A rating scale is an assessment instrument (technique) involving a set of categories designed to elicit information about a quantitative or qualitative attribute (based on pre-determined criteria). Through the use of the rating scale technique, the observer or rater categorizes the objects, events or persons on a continuum represented by a series of continuous concepts or numerals. There are four types of rating scale: nominal; ordinal; interval; and ratio.

Rating scales may be presented in six ways:

1. The graphic rating scale – various points are positioned along a line to form a continuum and the measurement is associated with its compared position along the line.
2. Numerical scale - A numerical scale is a rating scale that is used to measure or identify quantitative data.
3. Graphic scales
4. Percentage rating
5. Standard scales
6. Scales of cumulated points
7. Forced choice scales

3.40.3 Scale traceability

For metric scales the traceability problem is relatively simple -- all measurements have to be related to a single standard. For the qualitative scale types, the traceability problem can be more complicated because more standards are required.

Nominal scales typically have the greatest number of standards associated with them, usually one for each possible category on the scale. The standards may be descriptive or based on artefacts, such as standard reference materials.

Ordinal scales require a minimum of two standards, and in many cases require an approved or specified interpolating instrument.

Many interval scales can be expressed in terms of metric quantities, so the traceability problem is not too difficult. The log-ratio scale, for example, requires a definition of the multiplying constant, which can be defined without error, and a reference value, which in most cases takes the place of the unit on metric scales. All of the time scales (time of day, year, etc.) rely on measurements of time interval (a metric quantity) and an arbitrarily defined zero. Angle scales, such as latitude and longitude, also rely on angle interval and an arbitrary zero.

3.40.4 Evaluation through comparative and non-comparative scaling

In comparative scaling, evaluation involves comparing one thing of a certain type with another thing of the same type against a categorized set of criteria (e.g., one product with a specific function against another product with the same function). With noncomparative scaling, only one product is evaluated against a categorized set of criteria.

3.40.5 Degrees

The points along a scale may be referred to as degrees. A set of degrees (points) creates the scale. Therein, a 'property' of a scale is a 'degree'. A single degree within a scale represents:

1. A sub-delineation (subdivision) of a concept.
2. An iteration of a conceptual pattern.

A degree is a measurement of a whole concept (plane angle), defined so that the whole concept (full rotation) is expressed in the categorically (typologically) sub-divided manner of scales/levels of measurement (e.g., ordinal or interval; 360 degrees). Together, the categorical sub-divisions represent the whole set of possible divisions. For instance, there are 360 of the unit "degree" in the full rotation around a plane angle. A degree is a subdivided point along a conceptual line.

Here, the term 'range' refers to the bounds/endpoints of the concept or system. It represents the range of that which is possible, the range of possible numeric or conceptual values. For instance, the range of common subdivisions chosen for a plane angle (because history, or logic) is 360 (as, the count of 1,2,3,4,...). Hence, the range would be 0 at one boundary and 360 at the other (in a 2-axis/attribute system). Presently, the concept gender, and its relationship to genetic expression has a 2 attribute/characteristic [system]: female and male. In terms of gender, the scale [of measurement] has 2 attributes/degrees, which represent the entire range of possible attributes/degrees.

The range is the mapping of an attribute in the real world to a mathematical system. The mapping itself can be seen as a function behaving according to set of rules

NOTE: *Only information with a value (or number) assignment can be processed statistically.*

3.40.6 Statistics

Statistics is a type of mathematics. It is a mathematical data analysis system where uncertainty is fundamental, and the results are always expressed in terms of probabilities. Therein, models serve as both inputs and outputs of statistical analyses. Statistical analyses begin and end with models.

Statisticians often refer to the "levels of measurement" of a variable, a measure, or a scale to distinguish between measured variables that have different properties. In statistics, the term measurement is used more broadly and is more appropriately termed 'scales of measurement'. The term 'scales of measurement' refers to ways in which variables/numbers are defined and categorized. Each scale of measurement has certain properties which in turn determines the appropriateness for use of certain statistical analyses.

> **NOTE:** *Drafting scale rulers read architectural and engineering drawings.*

3.41 Level of measurement: Qualitative (categorical)

Qualitative information is determined by the nominal and ordinal level of measurements, which represent techniques (or scales).

3.41.1 Nominal [-level of measurement, scale] - categorized data, name, not numerical

Nominal levels of measurement are used to distinguish between features only on the basis of qualitative information. Nominal data does not imply quantitative differences. The only understanding conveyed is that two things (objects/events) of the same category have a difference (i.e., that in the category containing A and B, that A is different to B. It is meaningless to add, subtract, multiply, or divide nominal data. Attributes are only named; weakest. The assignment of a number. Classification of objects where the fact that the objects are different is preserved. Categorical data and numbers that are simply used as identifiers or names represent a nominal scale of measurement. This is not a 'scale' because it does not scale objects along any dimensions; it simply labels objects. Gender, for example, is a nominal scale: female = 1 & male = 2. The nominal type differentiates between items or subjects based only on their names or (meta-) categories and other qualitative classifications they belong to; thus dichotomous data involves the construction of classifications as well as the classification of items. Discovery of an exception to a classification can be viewed as progress. Numbers may be used to represent the variables but the numbers do not have numerical value or relationship: for example, a Globally unique identifier. Nominal scales were often called qualitative scales, and measurements made on qualitative scales were called qualitative data. However, the rise of qualitative research has made this usage confusing. The numbers in nominal measurement are assigned as labels and have no specific numerical value or meaning. No form of mathematical computation (+,-x etc.) may be performed on nominal measures. The nominal level is the lowest measurement level used from a statistical point of view.

Nominal is without order. Nominal could be considered a qualitative scale technique for grouping into unique categories (e.g., eye color).

Nominal is hardly measurement. It refers to quality more than quantity. A nominal level of measurement is simply a matter of distinguishing by name, e.g., 1 = male, 2 = female. Even though we are using the numbers 1 and 2, they do not denote quantity. The binary category of 0 and 1 used for computers is a nominal level of measurement. They are categories or classifications.

Nominal data refers to data which can be organised into categories e.g. gender: men and women, type of pet: cat, dog, fish, etc. Nominal data does not refer to numbers or quantities. You can't divide a dog by 2 (or at least you shouldn't).

1. A variable that has a nominal-level measurement scale is commonly referred to as a nominal-level variable, or simply, a nominal variable.
2. Define classes or categories, and then place each entity in a particular class or category, based on the value of the attribute.
3. The empirical relation system consists only of different classes; there is no notion of ordering among the classes.
4. Any distinct numbering or symbolic representation of the classes is an acceptable measure, but there is no notion of magnitude associated with the numbers or symbols.

> **TERMINOLOGY:** *Dichotomous – nominal, but two categories only (e.g., male/female).*

3.41.2 Ordinal [-level of measurement, scale] - ordered categorized data, semantic data, name with order, positional, categories with numerical order only

> **NOTE:** *This scale is qualitative, but seemingly quantitative.*

Ordinal scales involve differentiation by class, but they also differentiate within a class of features on the basis of rank according to some qualitative measure. Only rank is involved in ordinal scales. We are able to say that object A has a higher rank than object B, but we cannot say by how much.

Ordinal is nominal with order. Ordinal could be considered is a qualitative scale technique for grouping categories with order (e.g., mild, moderate, or severe; or, 1,2,3). In application, this can sometimes be difficult to separate from nominal.

> **APHORISM:** *Question: What is ordinal?*
> *Response: What number are you in line?*

Values assigned to the levels of a variable simply indicate that the levels are in order of magnitude.

Scale for ordering observations from low to high with any ties attributed to lack of measurement sensitivity (e.g., score from a questionnaire).

Ordinal refers to order in measurement. An ordinal scale indicates direction, in addition to providing nominal information. Low/Medium/High; or Faster/Slower are examples of ordinal levels of measurement. Ranking an experience as a "nine" on a scale of 1 to 10 tells us that it was higher than an experience ranked as a "six." Many psychological scales or inventories are at the ordinal level of measurement.

The scale is constructed so that there is an order to all divisions.

> **NOTE:** *Subjective measurement scales may otherwise be known as semantic measurement scales.*

Calculation is possible with ordinal information, through the use of numbers which may be applied to these qualitative scales, which when algebraically calculated, will give a mathematical result (a.k.a., "score").

Ordinal data refers to data which can be put into an order or ranked. Individual items can be organised by importance, general size or some arbitrary preference. Ordinal data ignores the exact degree of difference between individual ranked items. Attributes can be ordered/ranked.

Rank order of that which is being measured. Ordered categories of data. Objects are ranked/ordered based upon a criteria, but no information about the distance between the values is given. An ordinal scale of measurement represents an ordered series of relationships or rank order. The ordinal type allows for rank order (1st, 2nd, 3rd, etc.) by which data can be sorted, but still does not allow for relative degree of difference between them. Examples include, on one hand, dichotomous data with dichotomous (or dichotomized) values such as 'sick' vs. 'healthy' when measuring health, 'guilty' vs. 'not-guilty' when making judgments in courts, 'wrong/false' vs. 'right/true' when measuring truth value, and, on the other hand, non-dichotomous data consisting of a spectrum of values, such as 'completely agree', 'mostly agree', 'mostly disagree', 'completely disagree' when measuring opinion.

1. A variable that has an ordinal-level measurement scale is commonly referred to as an ordinal-level variable, or simply, an ordinal variable.
2. The empirical relation system consists of classes that are ordered with respect to the attribute.
3. Any mapping that preserves the ordering (that is, any monotonic function) is acceptable.
4. The numbers represent ranking only, so addition, subtraction, and other arithmetic operations have no meaning.

Common qualitative, ordinal measurement scales include the following scaling techniques:

1. Likert-type scale (qualitative)
    A. Evaluation-type: little, unsatisfactory, satisfactory, excellent.
    B. Frequency-type: never, rarely, occasionally, most of the tie.
    C. Agreement-type: strongly agree, agree, disagree, strongly disagree.
2. Semantic differential scale – scale includes semantic opposition.
    A. Slow <> Fast; Timely <> Untimely.
3. Summative scale

In this context a scale would give a score (qualitative), a dimension would be an actual measurement (quantitative). For example:

1. Dimension – Length of an object (quantitative).
2. Dimension - Percentage of organisms who die from the same mass dosage of a poison (quantitative; "lethal dose 100/50/30/10).

3.42 Level of measurement: Quantitative (cardinal, metric scale, numerical)

There are three quantitative levels of measurement note here:

1. Interval
2. Ratio
3. Absolute

3.42.1 Interval [-level of measurement, scale] – identified intervals, space between categories is identified

> **NOTE:** *Some statistics software packages may refer to cardinal and ratio data as 'scale'.*

Values assigned to the levels of a variable simply indicate that the levels are in order of magnitude in equal intervals. "Interval" itself means "space in between," which is the important thing to remember–interval scales not only tell us about order, but also about the value between each item. Interval scales provide information about order, and also possess equal intervals.

A degree represents the same underlying amount of that which is being measured (e.g., heat if temperature is measured), regardless of where it occurs on the scale.

Interval scales don't have a "true zero." For example, there is no such thing as "no temperature."

Interval scales are those that are known to be linear in some fundamental sense, and are the simplest scale type to allow meaningful comparison of differences. Interval scales typically have an arbitrary zero. Familiar examples include the latitude and longitude scales, which are used to determine position on the surface of the earth. The longitude scale requires two standards to define it: the position of the zero, which is arbitrarily chosen to be Greenwich, and the number of degrees in a full revolution of the earth, which is arbitrarily chosen to be 360. It is possible to compare changes in longitude meaningfully, or to add and subtract intervals of longitude, but it is still not meaningful to talk about ratios. Statements such as 'a country at 40 degrees of longitude is twice the country at 20 degrees of longitude' are nonsense.

Other examples of interval scales include all of the time scales that we use to tell the time of day, date and year, and the 4 mA to 20 mA current loop representation used by many industrial instruments (a symbol need not be a squiggle on paper). One of the earliest thermodynamic temperature scales, the centigrade scale, was an interval scale based on the definition of the melting and boiling points of water at 0 °C and 100 °C respectively. Because interval scales are the first that enable us to talk meaningfully about intervals, these are the first scales that allow us to do normal statistics, that is to calculate means and standard deviations.

Without a true zero, it is impossible to compute ratios.

Interval scales add information about the distance between ranks. To employ an interval scale we must use some kind of standard unit. For example, we differentiate between temperatures by using the standard unit of degrees celsius. We distinguish among elevations by using the arbitrary datum of mean sea level. We cannot multiply or divide interval scale data. For example, it would be incorrect to say 40 °C = 2 * 20 °C. Interval scales have no true or absolute zero. A temperature of 0 °C does not imply an absence of heat, it is just the point at which water freezes.

Cardinal data (also known as interval data) refers to data comprised of consistent units/intervals. Higher numbers mean more of something whereas lower numbers always mean less of something e.g. height, weight, time, temperature, etc. Cardinal data doesn't always have what's known as a 'true zero'.

Distance [between attributes] is meaningful. Differences between values are meaningful. Equal degree/rating (in a 'range') on a scale between two numbers (numerical values). A scale which represents quantity, and has equal units, but for which zero represents an additional point of measurement is an interval scale. The interval type allows for the degree of difference between items, but not the ratio between them. Examples include temperature with the Celsius scale, which has two defined points (the freezing and boiling point of water at specific conditions) and then separated into 100 intervals, date when measured from an arbitrary epoch (such as AD), percentage such as a percentage return on a stock, location in Cartesian coordinates, and direction measured in degrees from true or magnetic north. Ratios are not meaningful since 20 °C cannot be said to be "twice as hot" as 10 °C, nor can multiplication/division be carried out between any two dates directly. However, ratios of differences can be expressed; for example, one difference can be twice another. Interval type variables are sometimes also called "scaled variables", but the formal mathematical term is an affine space (in this case an affine line).

1. A variable that has an interval-level measurement scale is commonly referred to as an interval-level variable, or simply, an interval variable.
2. Interval scale carries information about the size of the intervals that separate the classes.
3. An interval scale preserves order, as with an ordinal scale.
4. An interval scale preserves differences, but not ratios. In other words, the difference between any two of the ordered classes in the range of the mapping is known, but computing the ratio of two classes in the range does not make sense.
5. Addition and subtraction are acceptable on the interval scale, but not multiplication and division.

NOTE: *Occasionally, in common parlance, the ratio scale is considered a second [categorical] class of cardinal measurement, with the first being the interval scale. In other words, there are two sub-categories of cardinal measurement: the interval scale and the ratio scale.*

3.42.2 Ratio [-level of measurement, scale] – measured intervals zero, relation to an absolute datum

Ratio data is the highest measurement scale. All forms of arithmetic operations can be meaningfully applied to ratio scale data. There is a meaningful "zero" value (a fixed origin), and ratios between values are meaningful. Equal degrees on a scale, and the non-existence of a degree is meaningful. The ratio scale of measurement is similar to the interval scale in that it also represents quantity and has equality of units. However, this scale also has an absolute zero (no numbers exist below the zero). Very often, physical measures will represent ratio data (for example, height and weight). If one is measuring the length of a piece of wood in centimeters, there is quantity, equal units, and that measure cannot go below zero centimeters. A negative length is not possible. The ratio type takes its name from the fact that measurement is the estimation of the ratio between a magnitude of a continuous quantity and a unit magnitude of the same kind (Michell, 1997, 1999). A ratio scale possesses a meaningful (unique and non-arbitrary) zero value. Most measurement in the physical sciences and engineering is done on ratio scales. Examples include mass, length, duration, plane angle, energy and electric charge. In contrast to interval scales, ratios are now meaningful because having a non-arbitrary zero point makes it meaningful to say, for example, that one object has "twice the length" of another (= is "twice as long").

In a ratio scale, the following is known:

1. The order
2. The exact value between units
3. An absolute zero
    A. A variable that has a ratio-level measurement scale is commonly referred to as a ratio-level variable, or simply, a ratio variable.
    B. The most common scale in the physical sciences.
    C. It is a measurement mapping that preserves ordering, preserves size of intervals between entities, and preserves ratios between entities.
    D. There is a zero element, representing a total lack of the attribute.
    E. The measurement mapping must start at zero and increase at equal intervals, known as units.
    F. All arithmetic can be meaningfully applied to the classes in the range of the mapping.

A scale in which the values assigned to the levels of a variable indicate both the order of magnitude and equal intervals, but in addition, assume a real zero. The real zero represents the complete absence of the trait that is being measured.

In addition to possessing the qualities of nominal, ordinal, and interval scales, a ratio scale has an absolute zero (a point where none of the quality being measured exists). Using a ratio scale permits comparisons such as being twice as high, or one-half as much. Reaction time (how long it takes to respond to a signal of some sort) uses a ratio scale of measurement -- time. Although an individual's reaction time is always greater than zero, we conceptualize a zero point in time, and can state that a response of 24 milliseconds is twice as fast as a response time of 48 milliseconds.

This type of scale is also known as the metric scale. Metric scales include all of the familiar SI scales of length, mass, thermodynamic temperature, etc. The mass scale is defined in terms of the prototype kilogram stored in a safe in a basement of the Bureau International des Poids et Mesures (BIPM) in Paris. All other measurements reported on the mass scale are expressed as ratios with respect to the kilogram. The standard used to define the scale is known as the metric or the unit of the scale. Metric scales are also known as ratio scales, and the literal translation of the word metrology, from the Greek metrología, is the study of ratios.

NOTE: *The measurement scales for counting oranges and apples are different, because they have different metrics, one orange and one apple respectively, and one cannot take one apple from two oranges and obtain a meaningful result*

The log-ratio scales form a special class of interval scales that are actually based on metric quantities. Because of the very large range of values encountered, it is often convenient to transform metric measurements to a logarithmic scale. These scales are typically constructed as value on log scale = constant × log (value/reference value).

There are two definitions required to define a log-ratio scale: the multiplying constant and the reference value. Examples of such scales include the various decibel scales, the visual magnitude of stars, and the Richter scale for the energy dissipated in earthquakes. On these scales equal intervals correspond to constant multiplying factors of the underlying metric quantity. An interval of 10 dB corresponds to a 10 times increase in power, five steps of visual magnitude correspond to 100 times decrease in the brightness of stars, and two steps on the Richter scale correspond to a 1000 times increase in the energy dissipated in an earthquake.

The progression of scales given above (from lower to higher mathematical operations) suggests that as the nature of quantities and measurements becomes well understood, the associated scales evolve towards metric scales (i.e., a scale with a natural zero).

Some scales can never be metric: colour will always be a three-dimensional scale based on two interval quantities and one metric quantity, and the Rockwell hardness scales will always be ordinal scales.

The counting/natural scale is a metric scale.

With metric scales, an additional possibility is available, namely geometric or harmonic analysis, which is based on distributions measured in terms of ratio rather than interval. An analysis of quantities measured on log-ratio scales using interval statistics is effectively a ratio analysis of the underlying metric quantity.

3.42.3 Absolute [-level of measurement, scale] – measured intervals with true zero, relation to an absolute datum

The ratio scale with an absolute zero (a.k.a., true zero) is sometimes called absolute. No transformation (other than identity) is meaningful. The non-existence of a degree means the non-existence of something in the real world.

An absolute scale is a system of measurement that begins at a minimum, or zero point, and progresses in only one direction. An absolute scale can only be applied to measurements in which a true minimum is known to exist.

Absolute scales are used when precise values are needed in comparison to a natural, unchanging zero point. Measurements of length, area and volume are inherently absolute, although measurements of distance are often based on an arbitrary starting point. Measurements of weight can be absolute, such as atomic weight, but more often they are measurements of the relationship between two masses, while measurements of speed are relative to an arbitrary reference frame.

Defined over a closed set (e.g., objective probability).

1. A variable that has a absolute-level measurement scale is commonly referred to as a absolute-level variable, or simply, a absolute variable.
2. The absolute scale measurement process involves counting the number of elements in an entity set (a conceptualization or pattern). In other words, the absolute scale measurement process involves quantizing the quantity of separations/patterns in a conceptualization or pattern (in a, concept pattern).
3. The measurement for an absolute scale is made by counting the number of elements in the entity set.
4. The attribute (measure) always takes the form:
   A. Number of occurrences of conceptual entity.
   B. Quantity, Unit.
5. There is only one possible measurement mapping.
6. All arithmetic analysis of the resulting count is meaningful.

4 The unit

A measurement necessarily involves a reference frame, and therefore, units. In decades past, there were numerous units, which had little in common with each other. The first coherent system of units only appeared with the French revolution, and it has been given the familiar name, the Metric System.

A unit of measurement (a.k.a., measurement unit, or just, unit) is a real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the second quantity to the first one as a number. In physics, a given unit carries the semantic information of a [physical] property [of the universe] and an associated quantity. Hence, a unit is a standard measurement of the same quantity, for the purpose of comparison.

SEDRIS provides useful definitions of terminology herein:

- *Environmental Data Coding Specification (EDCS): 3 Terms, definitions, symbols and abbreviated terms*. SEDRIS Standards. Accessed: January 7, 2020. [standards.sedris.org]
- *Environmental Data Coding Specification (EDCS): 2 Normative references*. SEDRIS Standards. Accessed: January 7, 2020. [standards.sedris.org]

NOTE: *More general, a unit signifies the presence of a quantity or quality.*

The value of a quantity is generally expressed as the product of a [measurement] number and a [measurement] unit. Instances of the same unit category may have a quantitative difference represented by the number. It could be said that a measurement unit is a standardized quantity of a physical property, used as a factor to express occurring quantities of that property. A 'unit of measurement' is a definite magnitude of a quantity (number), defined and applied, that is used as a standard for measurement of the same quantity. Any other value of that quantity can be expressed as a simple multiple of the unit of measurement (e.g., metric is multiples of ten, and time is multiples of 60). In other words, a measurement unit is a standard that is used to measure some physical quantity. A report of a quantitative measurement is not meaningful without its units.

A 'measurement unit' is also known as a:

1. Unit of measurement or units of measure
2. A measurement unit
3. A measurable unit
4. A unit

The measurement of a different sub-category of existence is signified by different units. Therein, units [of measure] are names (not numbers) that signify (characterize) the type of property (concept or system) under measurement, and they associate a standard of comparison to which each is related.

> **NOTE:** *The condition of a system described by its properties (temperature, pressure, etc ...) is defined as its 'state'. At a given 'state', all properties of a system have fixed values. These values may, or may not, fluctuate in time – they may have different values at different states (i.e., dynamic; they may have different values).*

In the language of measurement, quantities are quantifiable aspects of the world, such as time, distance, velocity, mass, temperature, energy, and weight, and units are used to describe their magnitude (or quantity).

A unit must be related to the aspect of the object (system or event) to be measured. For example, a unit of area is required to measure area; area cannot be measured without a linear tool. A linear tool measures length. "We" give a name to a set amount length (i.e., a meter is a meters length). "We" give the name 'meter' to the set quantity length of a 'meter'. Meter and length are both units. The unit name is 'meter', and the dimensional unit [quantity] is 'length'. Once a reference point is established (i.e., a meter is a meter in length), then a scale of units may be created using prefixes, such as kilo (1000, 10-3) and milli (-1000, 10-3). Hence, kilometer is 1000 (or 103) meters, and nanometer is -1000 (or 10-3) meters.

A unit of measurement is relatively arbitrary but consistent. Why is a meter [distance] the quantitative length it is? Why is a second of time the quantitative time it is? These dimensional quantitative amounts have been selected by humans for their ease of perceptual comprehension as quantitative unit reference points from which to base (i.e., compare) existence within a similar conceptual category (Read: dimension).

A measured or counted quantity has a numerical value (e.g., 2.48) representing the quantity, and a unit signifier (whatever there are 2.48 of) representing the constant (fundamental or derived). Hence, when communicating measurements, it is essential to write both the value (#) and the unit (linguistic signifier) of each quantity/constant.

Note that it should not necessarily be presumed that within a single human and/or computational organization where most measurements tend to be reported consistently in the same units, that users will correctly infer the units when they are omitted. The omission of units may lead to unnecessary errors (or, "mistakes"). Further, automatic conversion and computer-assisted calculations become easier with the proper use of unit. It is optimal to always indicate the units.

The first step to check the validity of an equation or an expression in physics is to look at the unit. The units involved in the construction of equations will influence (determine) the form of the equation.

Dimensions and units

In application, the term 'unit' means essentially the same thing as the term 'dimension'. The terms unit and dimension are often used synonymously. However, there are slight differences in definition and rule application in the two alike terms when dimensions are defined as the conceptual quantities of a system, and units represent the name(s) given to a set quantity (quantitative amount) of the dimension. Units and dimensions have the following definitional differences:

1. An equation in which the units balance on both sides of the equal sign is called **coherent**.
2. An equation in which the dimensions balance on both sides of the equal sign is called **homogeneous**.
3. A unit system constructed so that all physical laws are represented by coherent equations is called a **coherent unit system**. Physics, chemistry, and most of engineering are built upon coherent systems.

Units and dimensions have the following rule differences:

1. Dimensions combine by the ordinary rules of algebra. Units do also.
2. Terms which are added or subtracted must have the same dimensions and the same units.
3. Quantities on either side of the equal sign must have the same dimensions and the same units.
4. Powers are dimensionless and unitless (though factors within them may have dimensions and units).
5. Percentages are dimensionless quantities, since they are ratios of two quantities with the same dimensions, and hence, have the same units.
6. dy/dx and ∂y/∂x have the dimensions and the units of y/x (look at the formula for the definition of the derivative).
7. ∫ y dx has the dimensions and the units of yx.
8. Arguments of sin, cos, tan, log, etc., must be dimensionless, but may have units.
9. Sin, cos, tan, log, etc., are dimensionless and unitless.
10. The mathematical constants pi (π) and e are dimensionless and unitless. Specific gravity, being a ratio of two densities, is dimensionless. It has no unit name. Index of refraction, a ratio of two speeds (of light), also has no unit name. Pi (π), the ratio of a circle's circumference to its diameter is therefore dimensionless and has no unit name.
11. Sometimes measurables of physically different quantities have the same dimensions. The commonest example is work and torque: both result from multiplying force by distance. In these cases the unit names are often assigned in a distinctive manner. Names of work units, erg,

Newton, etc., are never used for torques.

12. It is also possible for different quantities with different unit names to have the same dimensions. The quantity, 'luminous flux', has the unit 'lumen'. A light's source strength is expressed in the unit 'candela'. A one candela source is said to emit 4π lumens. It may be written, 4πC, where C is the source strength and F is the flux. 4π is dimensionless, so C and F have the same dimensions, even though representing distinctly different quantities with different unit names.

4.1 *Measurable, countable, and Non-countable units*

Measurable units (as opposed to countable units) are specific values of dimensions [of a system] that have been defined for communication and construction, such as grams for mass, seconds for time, centimeters or feet for length, etc. 'Units of measure' are the unit names given to these concepts (dimensional quantities), including any prefixes and/or suffixes. For instance, the mile and kilometer are both 'units of measure' (recognized as 'unit names') for the 'dimensional quantity' of length. Similarly, pounds, kilograms, and grams are recognized 'units of measure' for mass.

There are also countable and uncountable units (from the two types of nouns in the context of measurement). Countable nouns, otherwise known as unit nouns, have a singular and plural form (e.g., chair, chairs; 1 chair, 2 chairs, 3 chairs, etc. Uncountable nouns are also called mass nouns. Mass nouns only have one form, and hence, are fixed as a singular or plural. Mass nouns are not countable. For instance, it is not possible to say 1 furniture, 2 furnitures, 3 furnitures, etc. Mass nouns cannot be used with an a/an. However, it is possible to use "a part of / an system of" for singular and "some" for plural. For instance, it is possible to say, "a piece of furniture" or "some furniture", and "an item of cloth" or "some cloth".

Non-countable mass nouns (e.g., water and air) can be made into countable nouns by adding units of measure, such as:

1. A glass of water.
2. A liter of air.
3. A cubic meter of land.

Hence, there are:

1. Measurable units (measurable nouns)
2. Countable units (unit nouns)
3. Non-/Un-countable units (mass nouns)
4. Non-countable (mass noun) + unit of measure (known measurable quantity) = Countable units

Whereas,

1. **Unit nouns** usually refer to one or many separate items or units. For instance:
   - A system, systems; a habitat, habitats; an ecology, ecologies; a length, lengths, a width, widths; a mass, masses; a unit, units.
2. **Mass nouns** usually refer to:
   A. Ideas, concepts (e.g. approval, employment)
   B. Substances (e.g. meat, metal)
   C. Liquids (e.g. water, beer)
   D. Powders (e.g. dust, sugar)

There are 4 types of mass nouns:

1. Singular nouns that are always mass nouns (no plural form).
2. Plural nouns that are always mass nouns (no singular form).
3. Nouns that can be unit nouns or mass nouns and have the same meaning.
4. Nouns that can be unit nouns or mass nouns but have different meanings.

In many languages, 'information' is a unit noun, and has both singular and plural forms (information and informations respectively). However, in English and the community in general, 'information' is a singular mass noun; there is only information, and no informations.

4.2 *Unit taxonomy*

The concept of a 'unit' is taxonomically classified according to the three components of classification, identification, and nomenclature (naming):

4.2.1 Classification

Unit classification includes:

1. **Concept naming** - the logical, orderly naming of units based on derivation location.
   A. **Basic or fundamental units (a.k.a., base units, fundamental units, and dimensional units, fundamental dimensions)** - The smallest set of quantities that are accepted by definition. The basic measurables (a.k.a., fundamental measurable units and dimensions) are the basic/fundamental measurable units or dimensional units of a system. A base unit (also referred to as a fundamental unit) is a unit adopted for measurement of a base quantity (an axiomatic physical, natural property of existence or reality, fundamental dimension). A base quantity is one of a conventionally chosen subset of physical quantities, where no subset quantity can be expressed in terms of the others. The basic measurables of a system are called the "dimensions of the system".

Practically, they are the dimensions (parts/coordinates separated axiomatically) of the system. Here, the use of the word, "dimension", is analogous to its use in analytic geometry. In space, any point can be specified by its coordinates measured along axes of a three-dimensional coordinate system (generally signified in a standard manner as: x, y, and z). The dimensions of a quantity do not have an inherent unit association. The dimensions, and hence, units, arise from the logical interrelations between quantities, reflecting the structure of physical laws and definitions.

B. **Dimensional units** – concepts identifying the base/fundamental dimensions of a system.

C. **Base/fundamental units (a.k.a, metric units)** – the dimensional quantity associated name. For example, the metre unit as length dimension, kilogram unit as mass dimension, second unit as time dimension, kelvin unit as temperature dimension, etc.

D. **Multiple units** - express (by name) multiples or fractions of base units, such as minutes, hours, and milliseconds (for time), all of which are defined in terms of the base unit of time, a second. Multiple units are defined for convenience rather than necessity: it is simply more convenient to refer to 3 years than to 94608000 seconds.

E. **Derived units (sub-units)** - recognized by the dimensions and can be defined as the complete algebraic formula for the derived unit. In a system of measurable units, any derived measurable will be expressed as an algebraic combination of the basic/fundamental measurables (dimensions or basic units) of the system. Derived units are based on base/fundamental units, and can always be represented by these units. In other words, derived units are composed of several other units combined together.

2. **Concept Modifying** - modifying the name of a unit to indicate scale.
   A. **Unit [multiple] prefixes**
   B. A quantity of a unit can be re-written using a different logical name via a prefix multiplication scale. A prefix precedes the associated unit symbol to form a multiple or sub-multiple. This scale re-framing may make reading and calculation of the data more efficient (if human) or less efficient (if computer). For a human, it is easy to multiply by 10, for instance. Metric [unit] prefixes include: deca-, hector-, kilo-, deci-, centi-, milli-, etc. In total, there are twenty prefixes that have been officially adopted to be used with the Metric Unit System.

4.2.2 Identification

Procedures and methods for determination of an unknown unit.

4.2.3 Nomenclature

The logical naming of all the units in the taxonomy. Note here that there is little logical linguistic naming between the multiple units for the base unit of time, the second and a multiple unit of time, the minute and the hour. What is the relationship between the letters that compose the dimension 'time', the base unit 'second', and the unit multiples minute and hour? The words minute and hour do not appear logically related to 'second' or 'time'.

Summarily, classification refers to the sub-organization of unit-type concepts. Identification ensures that it is possible to procedurally (methodically) determine the unit for a known quantity. And, nomenclature ensures that names are logical, and hence, easy to recall and use.

> **NOTE:** *A 'dimensional analysis' is a scientific analyses conducted to determine the basic/fundamental measurable (measurable units) of a system.*

4.3 Fundamental and derived units [of measurement]

The basis of the physical sciences is a set of names, definitions, and equations, which allow for awareness, experimentation, and adaptation to a physical environment (i.e., our physical reality).

Not all quantities require a unit of their own. Using physical laws, units of quantities can be expressed as combinations of units of other quantities. Only a small set of units is required from which a more complex functional set can be built. The small set of required units of physical quantity are called base units (a.k.a., fundamental units), and all others units are derived [units]. Derived units are a matter of convenience, as they can always be expressed in terms of basic units.

> **CLARIFICATION:** *A [physical] quantity is a quantifiable [physical] aspect/attribute of the world (the universe, nature, reality), such as time, distance, velocity, mass, temperature, energy, and weight. A 'physical quantity' is a characteristic (property or quality) that can be measured, and which follows the laws of physics (which, describe and/or predict behavior and relationships). Here, physical quantity units are used to describe the magnitude or quantity of a physical aspect/attribute of the world.*

A base quantity is characterized by the following two principles:

1. Base quantities are those quantities which are distinct in nature and cannot be expressed in the form of other quantities.
2. Base quantities are those quantities on the basis of which other quantities can be expressed.

Similarly, a ‘fundamental measurement’ is characterized by the following two principles:

1. Measurement that is not derived from other measurements.
2. Measurement that is produced by an additive (or equivalent) measurement operation.

Hence, the two types of measurable physical quantities (i.e., physical units of measure) are:

1. **Axiomatic (base/fundamental quantity)**: A quantity that cannot be expressed in terms of other quantities. A quantity that is axiomatic (i.e., fundamental or base), and hence, cannot be defined in terms of the others. Those few which cannot be defined in terms of others, the “basic/fundamental measurable or dimensions”, are defined through operational definitions (by specifying a measurement process).
2. **Derived (quantity)**: A quantity that can be expressed in terms of other quantities (/units). This type is not axiomatic, and is defined (“derived”) algebraically in terms of other quantities.

In other words, the two physical units of measurement are:

1. **Fundamental units (a.k.a., basic units, fundamental measurable, basic measurable, and dimensions [of a system]** - Those defined by specifying a measurement process (i.e., by operational definitions). A base unit (also referred to as a fundamental unit) is a unit adopted for measurement of a base (fundamental axiomatic) quantity/constant. These are so directly connected with measurement that they are defined by the measurement process. Fundamental units describe axiomatic existent quantities (given what is known) from which all other units can be derived.
   - Base quantities are those quantities that are common to any object.
2. **Derived units (a.k.a., defined measurable)** - Those defined by algebraic mathematical equations in terms of other previously defined and/or fundamental measurables (measurable units).
   - Derived quantities are quantities formed by combining two or more base quantities (using multiplication or division).

Fundamental units may be perceived of from several different problem-oriented contexts:

1. In <u>mechanical problems</u>, a fundamental set of units is mass (M), length (L), time (T). With this fundamental system, velocity V = LT −1 and force F = MLT −2 are derived units. Alternatively, if instead, force (F), length (L), and time (T) are the fundamental system of units, then mass M = F L−1T 2 is a derived unit.
2. In <u>thermodynamic problems</u> (i.e., problems involving heat flow), the concept of temperature (measured, for example, in Kelvin) is a fundamental unit.
3. In <u>problems involving electromagnetism</u>, current is introduced as a fundamental unit (measured, for example, in Amperes in the SI system) or charge (measured, for example, in electrostatic units in the cgs system).
4. In <u>problems involving relativistic mechanics</u>, if mass (M), length (L), and time (T) are fundamental units, then the speed of light c is a dimensional constant (c = 3 × 108 ms−1 in SI-units). Therein, ‘c’ may be set to equal 1 (c = 1), and mass (M) and time (T) are the fundamental units. This means that length is measured in terms of the travel-time of light (one nanosecond being a convenient choice for everyday lengths).

> **NOTE:** *The ‘Rasch Analysis’ operationalizes ‘fundamental measurement’ based on ordered qualitative observations. Therein, ‘voltage’ as charge pressure, no matter the scale, is the most fundamental measurement for energy.*

4.4 The fundamental, base physical dimensional units [of measurement]

Today, there are seven scientifically recognized basic (base, fundamental) units of measurement, as that which is perceived as a fundamentally constant unit/quantity in the universe. Every other perceived [unitized] measure[able] is derived from those seven.

In other words, in physics, there are seven fundamentally perceivable:

1. Base quantities (detailed in the International System of Quantities, ISQ).
2. Fundamental dimensions of an axiomatic physical existence.
3. Units of [physical] measurement.

In physics, there are seven defined and measurable (dimensional) units. However, all seven fundamental units can be derived with three-four of the fundamental units.

1. Mechanics requires four fundamental measurable unit dimensions:
   A. Kinematical (3 units)
      1. Mass (kilograms)
      2. Length or distance (meters)
      3. 3 length dimension (x,y,z)
      4. Time (seconds)
   B. Electrical (1 unit)
      1. Ampere
2. It could be said that our human experience of the world encompasses five dimensions:
   A. Three linear spatial dimensions (x,y,z).
   B. One mass dimension.
   C. One temporal dimension (time).
3. Electricity requires two fundamental measurable units:
   A. Voltage (eV or Volts, depending on scale)
   B. Time (seconds)

However, it could be said that time is the only true unit of measure; because without time, no change can occur, and thus, no measurement. Measurement involves the perception of a change from a baseline, and change cannot occur without time. Time gives everything its existence, but it is not the true unit of measure. The nature of time is to flow (iterate), and the nature of consciousness is to experience the rate. Time (iteration) is essential in measurement in principle, for instance:

1. A mole is an exact number of "atoms". Measuring a mole requires time.
2. A candela is a measure of "luminocity". Luminocity is dictated by wavelength and frequency. A wave length has no length without time to travel said length, and frequency cannot be determined without measuring this length travelled over a time.
3. A meter is a measure of "length". The very action of measuring length requires time. Can you measure this without time? Grab a 'tape measure' and try. You have already failed as taking the measurement takes time.

That which involves a system in time, involves:

1. **Duration** (of time of system)
2. **Volume** (3x length dimensions of system in time)
3. **Concentration** (mass of system in time)
4. **Intensity** (electric current of system in time)

Take note here that the common properties of physical systems include, but are not limited to:

1. Pressure (P)
2. Temperature (T)
3. Volume (V)
4. Density (D)
5. Mass (M)
6. Energy (E)

It is common in the realm of the elementary particle physics to redefine units so that speed of light and Plank's constant become equal to one, c=1 and h =1. This imposes two constraints on the three kinematical units, and therefore, provides a choice one of the three kinematical units. The units of electrical charge, also, can be, and are redefined (see below). Such system of units is often referred to as Natural Units (natural for the elementary particle physics, that is). The kinematical unit of the choice is energy, E, and it is usually measured in eV (keV, MeV, GeV, TeV). Once c and h are fixed (c=1 and h =1), all other kinematical units can now be expressed in terms of units of energy.

> **INSIGHT:** *If every point in the universe (i.e., every proton has the information of all other protons in the universe) has all the information about the universe (a holographic system), then the universe has the ability to self-organize. Then, every point knows exactly how to self-organize, because all the information is present in every point. We are feeding the universe information, and the universe is feeding us information through all the protons we are made of.*

4.4.1 Geometrized units [of measurement]

A geometrized unit system or geometric unit system is a system of natural units in which the base physical units are chosen so that the speed of light in vacuum, c, and the gravitational constant, G, are set equal to unity.

- c = 1
- G = 1

The geometrized unit system is not a completely defined or unique system: latitude is left to also set other constants to unity. We may, for example, also set Coulomb's constant, ke, and the electric charge, e, to unity.

- ke = 1
- e = 1

The reduced Planck constant, ħ, is not equal to 1 in this system (Stoney units), in contrast to Planck units.

4.5 Physical constant, natural units [of measurement]

In physics, **natural units** are physical units of measurement based only on universal physical constants (a.k.a., the fundamental constants of physics; invariant quantities), and not on human constructs. There are many physical constants in science.

For example,

1. The elementary charge ‘e’ is given as the natural unit of electric charge, and
2. The speed of light ‘c’ is given as the natural unit of speed.

A physical constant (a.k.a., fundamental physical constant) is a physical quantity (a.k.a., fundamental physical quantity) that is understood to be both universal in nature and having constant value in time. It is contrasted with a mathematical constant, which has a fixed numerical value, but does not directly involve any physical measurement. Natural units are “natural” because the origin of their definition comes only from properties of nature and not from any human construct (i.e., they can be experimentally demonstrated).

> **NOTE:** *Using dimensional analysis, it is possible to combine dimensional universal physical constants to define a system of units of measurement that has no reference to any human construct.*

Properties of the universe that are likely to have quantity may be represented as natural units. Natural units are intended to simplify particular algebraic expressions appearing in the laws of physics, or to normalize some chosen physical quantities that are properties of universal elementary particles, and are reasonably understood to be constant.

The value of any one of these seven constants is written as the product of a numerical coefficient and a unit, Q = {Q} [Q], where Q denotes the value of the constant and {Q} its numerical value when expressed in the unit [Q]. By fixing the exact numerical value — that is, not assigning any uncertainty to it — the unit becomes defined, as the product of the numerical value and the unit must equal the value of the constant, which is invariant.

There are many natural units (as defined constants), including but not limited to:

1. The speed of light in vacuum - c
2. The Planck constant - h
3. The elementary [electric] charge - e
4. The Boltzmann constant - kB
5. The Avogadro constant - NA
6. The luminous efficacy - Kcd
7. The gravitational constant – G
8. The electron rest mass – me
9. The Josephson constant – KJ
10. The frequency of the ground-state hyperfine splitting of the caesium-133 atom - Δν(133Cs)hfs

A purely natural system of units has all of its units defined by physical constants. Usually, the numerical values of the selected physical constants defined in terms of these units are exactly dimensionless (1).

These constants should not be omitted from mathematical expressions of physical laws; though omission has the apparent advantage of simplicity, it may entail a loss of clarity due to the loss of information, which is otherwise is required for dimensional analysis. Omission of the constant precludes the easy cognitive interpretation of an expression in terms of fundamental physical constants, such e and c.

Throughout all of the formulations of the basic theories of physics, and their application to the real world, there appear certain fundamental invariant quantities. These categorical delineations in our understanding of the reality system of our experiences are called, [fundamental] physical quantities/constants (i.e., fundamental physical quantities and fundamental physical constants). These constants/quantities have specific and universally used symbols.

It is important to understand that most measurements are relative by nature, so only measurements (and units) as the basis for other measurements need to be solitary by nature. Those measurements (or units) that are presently understood to foundation all others are: time, current-voltage, mass, and length. All other measurements (i.e., all other measurables) are based on those units. For instance, velocity is distance per unit of time, Hertz is the number of voltage cycles per unit of time, and calories is the chemical energy (measurable as eV) released per unit mass, etc.

4.6 A unit system (system of units)

A system of measurement is a collection of measurement units, for various concepts of “measure” (i.e., dimensions and units; e.g. length, mass, time), where various units are mutually consistent, and interrelate in a standardized way. Practically, a system of units (a.k.a., unit system) forms a group of pre-determined reference amounts with logical naming. Simply, a standard[ized] set of units is called a ‘unit system’. In order to take (i.e., “make”) a quantitative measurement, a system of units is required; that is, a set of magnitudes with which to compare those things (properties/attributes) for which comparison (i.e., measurement) is desired.

A system of units is a necessary input for cooperatively measuring ourselves and our environment (Read: the cosmos, the universe, nature, reality).

A set of fundamental/basic units is otherwise known as a ‘system of units’. Different fundamental/basic systems of units are based on different choices of base units. A [basic, fundamental] system of units is a set of independent (axiomatic) units from which all other units in the system can be derived. The choice of fundamental units in a particular class of problems is not unique, but, given a fundamental system of units, any other derived unit may be constructed uniquely as a product of powers of the fundamental units.

4.7 Coherent versus incoherent unit systems

A unit system may be either coherent or incoherent. In order to establish a coherent system of units (e.g., the SI) it is necessary to first establish a system of quantities, including a set of equations defining the relations between those quantities. Units therein are consistently constructed and consistently named. Incoherent unit systems have units with no direct relation to each other, and when there are relations, they lack consistency (e.g., the Imperial and US systems). The units within incoherent unit systems are therefore difficult to remember and less efficient to work with.

A coherent unit system is built by choosing appropriately sized basic/fundamental dimensional units for the users' cognition and/or computational parameters. For instance, a meter is given the quantity it has been given, in part, because it is easy for human cognition. The units of other measurable dimensions will then be determined by their defining equations, as combinations of the units of the base/fundamental measurable dimensions, in the same manner as dimensions are determined.

A coherent unit system is a set of coherent axiomatic (base, fundamental) dimensional units that can be used to accurately understand and construct that which is conceptualized. The most widely used system of units is the International System of Units, or SI. There are seven SI base units; all other SI units (non-base) can be derived from these base units.

> **NOTE:** *A logically standardized system of units allows for efficiency in measurement, and hence, efficiency in design, development, and operation [of service systems].*

Take note that in physics, coherent unit systems can presently be built upon a set of basic units that includes only one of the following:

1. When mass is included within the set of basic units, the system is called '**absolute**'.
2. When force is included within the set of basic units, the system is called '**gravitational**'. The fps system is characterized by a gravitational unit of force, called the pound-force (lbf). The unit is so defined that a standard gravitational field exerts a force of one pound on a mass of one avoirdupois pound.

4.8 Common unit systems in use on the planet today

The four most common unit systems in use today are:

1. **The International Metric System (MKS, the [Decimal] Metric System)** is an absolute system. Its basic units are the meter, kilogram, and second. There are several variants of the metric system, including:
   A. **The International System of Units (SI, for System, International**) is the modern/revised form of the metric system, and is the most widely used system of measurement. It has seven basic units, including the meter, kilogram, and second.
   B. **The CGS (centimeter-gram-second) system** was once standard in physics.
2. **The FPS (foot-pound-second) system** was once standard in engineering, and is a gravitational system of units.
   - The FPS is an incoherent system.

These three systems (MKS-SI, CGS, and FPS) are all mutually coherent for most branches of physics, especially mechanics (but not including electricity and magnetism). In mechanics, the equations have the same form in all three. In electromagnetics the International System of Units (SI) is used; the FPS system does not account for electromagnetics.

Take note that in the Metric/SI systems, each different kind (dimension) of measurement has a root name, from which other names may be constructed by combining the name with a metric prefix. For instance:

1. Meter, a length unit, forms millimeter, centimeter, and kilometer.
2. Gram, a mass unit, forms milligram, centigram, and kilogram.
3. Liter, a unit of capacity (volume), forms milliliter, centiliter, and kiloliter.

> **CLARIFICATION:** *In mathematics, a 'metric space' is a set for which distances between all members of the set are defined. Those distances, taken together, are called a metric on the set. Therein, a metric space induces topological properties like open and closed sets, which lead to the study of more abstract topological spaces.*

4.9 The International System of Units (SI)

> **HISTORICAL NOTE:** *The metric system of measurement was developed during the French Revolution and was first promoted in the U.S. by Thomas Jefferson. Its use was legalized in the U.S. in 1866. In 1902, proposed congressional legislation requiring the U.S. Government to use the metric system exclusively was defeated by a single vote. As of 2017, outside of the several States (including The United States and Great Britain), there is almost no need to convert metric units into something else, because they use metric units as their physical measurement system. In the United States and Great Britain, multiple measurement systems are used, which*

introduces the potential for confusion and error, and leads to an inefficient use of time and effort (due to the added necessity to convert).

The International System of Units is generally seen written as either:

1. International System of Units (SI for Systeme international d'unites), or
2. International System of Quantities (ISQ)

The International System of Units (SI) is the most up-to-date version of the Metric System, and it is formalized as a State agreement that specifies a set of seven base (physical-quantity measurement) units from which all other State agreed upon units of measurement are formed.

The International System of Quantities (ISQ) is a system based on seven base quantities: length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity. Other quantities such as area, pressure, and electrical resistance are derived from these base quantities by clear non-contradictory equations. The ISQ defines the quantities that are measured with the SI units. The ISQ is defined in the international standard ISO/IEC 80000, and was finalised in 2009 with the publication of ISO 80000-1.

The definitions of the terms "quantity", "unit", "dimension" etc., that are used in the SI Brochure are those given in the International vocabulary of metrology, a publication produced by the Joint Committee for Guides in Metrology (JCGM), a working group consisting of eight international standards organisations under the chairmanship of the director of the BIPM. The quantities and equations that define the SI units are now referred to as the International System of Quantities (ISQ), and are set out in the International Standard ISO/IEC 80000 Quantities and Units.

The Metric System (a.k.a., SI) provides a logical and interconnected framework for all physical measurements. The International System of Units (SI) is a modernized, State constructed, version of the Metric System established by international State/corporate agreement.

The SI unit system includes two types of units based on derivation location (axiomatic or sub-derived):

1. SI Base (Fundamental, Metric) Units
   - Currently there are 7.
2. SI Derived Units

The SI includes a coherent set of unit prefix multipliers.

1. Metric prefixes (prefix multipliers)
2. Currently there are +/- 24.
3. What about prefixes for other multiples, such as 104, 105, 10-4, and 10-5? The prefix myria- (my-) was formerly used for 104, but it is now considered obsolete and it is not accepted in the SI. Apparently, no prefixes were ever accepted generally for 105, 10-4, or 10-5, or others.
4. Table: Metric prefixes: Prefix, symbol, meaning, exponential notation.

4.9.1 The 2018 Update to the International System of Units

Updates to 2018 International System of Units

1. Ampere - e is the elementary charge (which defines an ampere). The unit used to measure electrical current. An ampere is the current that, when flowing through two infinitely long, infinitely thin wires that are placed exactly 1 metre apart, would produce a certain amount of force. But infinitely long and thin wires are impossible to produce, so no one can actually test precisely what that value should be. Under the new proposal, an ampere will basically be defined based on the electrical charge of the electron and the proton - something that scientists will actually be able to measure.
   A. The Ampere is the only electrical unit among the seven SI base units. Hence, one might logically expect that all other electrical units, including the volt and the ohm, will be derived from it. But that is not the case. In fact, the only practical way to realize the ampere to a suitable accuracy now is by measuring the nominally "derived" volt and ohm using quantum electrical standards and then calculating the ampere from those values.
   B. In 2018, however, the ampere is slated to be re-defined in terms of a fundamental invariant of nature: the elementary electrical charge (e). Direct ampere metrology will thus become a matter of counting the transit of individual electrons over time.
   C. One promising way to do so is with a nanoscale technique called single-electron transport (SET) pumping. It involves applying a gate voltage that prompts one electron from a source to tunnel across a high-resistance junction barrier and onto an "island" made from a microscopic quantum dot. The presence of this single extra electron on the dot electrically blocks any other electron from tunneling across until a gate voltage induces the first electron to move off the island, through another barrier, and into a drain. When the voltage returns to its initial value, another electron is allowed to tunnel onto the island; repeating this cycle generates a steady, measurable current of single electrons.

D. There can be multiple islands in a very small space. The distance from source to drain is a few micrometers, and the electron channels are a few tens of nanometers wide and 200 nm to 300 nm long. And the energies involved are so tiny that that device has to be cooled to about 10 millikelvin in order to control and detect them reliably. (Stewart, 2016)
E. The ampere [A], is the unit of electric current; its magnitude is set by fixing the numerical value of the elementary charge to be equal to exactly 1.602 .10(-19) when it is expressed in the unit [As], which is equal to C. Thus we have the exact relation e = 1.602 .10(-19) [C]. The effect of this definition is that the ampere is the electric current corresponding to the flow of 6.242 10(18) elementary charges per second. The following is not true in SI: The present basic unit of electric current Ampere can't be basic unit because is defined with Coulomb and second. Ampere is not unique unit, because depends on other units. From this is obviously that the Coulomb has no relation whit any other units and because of that it's most convenient this unit to be proposed as basic units.
F. Previous to 2018: The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2 x 10–7 newton per meter of length.

2. Kelvin - redefined by linking it to the exact value of the Boltzmann constant. Previously, the Kelvin was defined as the triple point of water - the obscure point where water coexists as a liquid, gas, and solid. kB is the Boltzmann constant (which defines a kelvin).
3. Mole - NA is the Avogadro constant (which will define a mole). It was redefined in 2018 by linking it to the exact value of the Avogadro constant.
4. Meter - C is the speed of light (which will define a metre.
5. Second - ΔνCs is the tick of a caesium atom clock (which defines a second).
6. Kilogram - h is the Planck constant (which will define a kilogram). The Planck constant is measured by placing a known mass on one end of a scale, and then, counterbalancing it by sending an electric current through a movable coil of wire suspended in a magnetic field. The electromagnetic force therein is used to measure Planck's constant down to an accuracy of 34 parts per billion.
7. Candela - Kcd is luminous efficacy (which will define a candela).

A dimension is a property that can be measured (e.g., length, time, mass, or temperature) or calculated by multiplying or dividing other dimensions (e.g., length/time = velocity, length3 = volume, or mass/length3 = density).

4.10 Systems of units (in use today)

There are multiple systems of units, some of which are more intuitive and logical (depending on context), than others. There are four categories of unit [measurement] systems:

1. **Traditional unit systems** - Historically many of the systems of measurement which had been in use were to some extent based on the dimensions of the human body. As a result, units of measure could vary not only from location to location, but from person to person. Based on "arbitrary" unit values.
   A. **The Imperial System of Units** used in the United Kingdom and former colonies. It bases its measures on human anatomy (generally, the body parts of royals/imperials) and on common objects that humans use. Early on in human development, people used signifiers like body parts as their units of measurement. For instance, the imperial measurement unit known as a "foot" is about the length of a human foot. An "inch" is about the length of a human thumb. A cup is about the weight of a cup of water. A pound is about the weight of 2 cups of water. Note that there is no common base in the Imperial System. Conversely, in the Metric System, the base unit for all measurements is the number (the de-lineation) 10. The Imperial System includes, but is not limited to the following measurement sub-system units: the Foot, the Pound, the Gallon, and the Mile (statute mile and nautical mile).
   B. **The Market System of Units** used in the State of China.
   C. **The United States Customary Unit System** used in the United States
2. **The [Decimal] Metric Unit System** - A number of metric systems of units have evolved since the adoption of the original metric system in France in 1791. The current international standard metric system is the International System of Units (abbreviated to SI). An important feature of modern systems is standardization. Each unit has a universally recognized size. Both the imperial units and US customary units derive from earlier English units. Imperial units were mostly used in

the British Commonwealth and the former British Empire. Based on “arbitrary” unit values, formalized by standards.

A. **The [Decimal] Metric System of Units** (used for globally-coordinated projects): In the [decimal] metric system, every measure is a factor of 10 units from others. The metric (a.k.a., decimal metric) system uses base 10 for everything, which allows for easy calculation and scaling. It is, in terms of human-mind calculation, easier to work with direct powers of 10 proportions, than any other proportion, and particularly when units are of dissimilar proportion (e.g., inch, foot, yard, pound, ounce, etc.) and dissimilar by geographic location (e.g., UK gallon versus the US gallon). In other words, millimetres, centimetres, and kilometres are interchangeable, whereas feet, yards, and miles are not. When performing a mental calculation, someone doesn’t have to have in mind a table referencing how many inches are in a foot, how many ounces are in a pound.

B. In distance (reference is ‘meter’): 1M = 100centiM = 1000milliM = 0.001kiloM (or 0,001kiloM).

C. In volume (reference is ‘litre’ or ‘liter’): 10deciL = 1L = 1000milliL. 10cm x 10cm cube of water weighs about a kilogram, and is otherwise known as a litre.

3. **The Natural Unit Systems** – Unit values that have logically deduced or experimentally demonstrated to occur naturally in science.
   A. **Atomic units (au)** – system of units of measured used in atomic physics.
4. **State/legal Weights and Measures** - To reduce the incidence of retail fraud, many national statutes have standard definitions of weights and measures that may be used (hence “statute measure”), and these are verified by legal State representatives.
   A. **Units of currency** - A unit of measurement that applies to money is called a ‘unit of account’ in economics and ‘unit of measure’ in accounting. This is normally a currency issued by a State, or a faction thereof.

4.11 Systems of units used by Community

A ‘measurement unit system’ is a group of related measurement units.

Currently, there are two active measurement systems:

- Metric (Absolute) Measurement System
  - Contains Metric measurement units: Micrometer, Millimeter, Centimeter, Decimeter, Meter and Kilometer.
- Graphics (Imaging, Visualization) Measurement System
  - Contains GDI measurement units: Pixel, Point, Display, Document, Inch and Millimeter. In graphics measurement units are typically used to express the length, size or location of objects (i.e. some object dimensions).

When multiple unit systems are in common use, it is often necessary to convert the magnitudes of quantities from one system to another. This is accomplished by using conversion factors. Only the defined conversion factors for the base units are required since conversion factors for all other units can be calculated from them. Conversion factors are necessary for interconversion (Read: conversion between systems).

4.11.1 Measurement device units

Currently there are four types of measurement device units:

1. Device - device measurement units are the units of measure of the output device. For instance, with a computer display system, there is only one device measurement unit, and it is called, ‘pixel’.
2. Absolute - absolute measurement units are units that do not depend on the device. For example: inches and meters are absolute units, because their length does not depend on the output device (as in, the LCD display, which understands only pixels).
3. Relative - relative measurement units depend on the size of “something else”. In a system, relative measurement units are those units that depend on the size of the parent or root objects containing the object.
4. And also, physical and abstract measurement units

4.12 Unit conversion

> **NOTE:** *Units of measurement are not ratios, but ratios are necessary to convert between one unit of measurement and another.*

It is possible to have units that may be converted within and between unit systems. Conversion within a system may be either:

1. *Between unit prefixes* (e.g., between milliseconds and nanoseconds for the dimensional unit ‘time’)
2. *Between units* (unit quantities) with the same dimension (e.g., between seconds and hours for the dimension ‘time’)

Conversion between systems is:

- Between units (unit quantities) with the same dimension (e.g., between feet in FPS and meters in IS for the dimension 'length')

Two units (as in, unit names) measuring the same thing, but from different systems, are referred to as equivalents. If a task works in one unit system, but requires input from another unit system, then equivalent units for the specific issue, from the other unit system, must be identified. For instance, if a task uses the imperial unit system, but a specific sub-task requires a metric measurement, then the two systems can be converted between once a metric equivalent (i.e., equivalent metric unit) is identified and its conversion factor is determined.

- **Conversion factors** are homogeneous, but may be incoherent. Their primary use is to transform equations from one coherent unit set to another.

Unit conversion is the process (technique) of exchanging one unit of measure for another unit of measure, while maintaining the associated value (or count).

It is possible to convert within and between unit systems wherever the units mean (measure) the same dimension, object, or event ("thing").

One way to avoid an additional conversion task (and hence, conversion formulas) is to design and apply a single, coherent, and updatable measurement system, such as, the metric system.

Therein, one way to avoid an additional conversion task (and hence, conversion formulas) is to not use unit scale [multiplication] prefixes. However, not doing this can make reading and calculation challenging for humans.

Hence, it is possible (given similar conception) to convert into and out of any other system, and between different levels-of-scales within a single [measurement] system.

4.12.1 Between unit [scale] prefixes

Converting [a quantity] within the metric system [to a different level-of-scale] is known as 'metric conversion' (i.e., intra-metric unit multiplier conversion vs. inter-metric unit multiplier conversion between different measurement systems).

In the metric system, conversion occurs between multiplication prefixes, which include, but are not limited to: kilo, mega, giga, milli, micro, and nano.

Scientific notation is: M x 10n

- M is the coefficient 1<M<10
- 10 is the base
- N is the exponent or power of 10

4.12.2 Unit commensurability and incommensurability

During a task that involves a non-unified unit system, there may be unit types (with unique names) that measure the same thing (same concept), but are based on different [reference] standards. One unit either has a common basis of measured meaning with another (or others), or it does not. If a single unit is present, then commensurability is not an issue.

1. **Commensurable units** have a common basis [for the transfer] of a set value. In order to transfer, a ratio that equals 1 must be present.
2. **Incommensurable units** do not have a common basis [for the transfer] of a set value, and hence, a set value (quantity) of that unit cannot be transferred.

The term 'incommensurable' means 'no common measure', having its origins in Ancient Greek mathematics, where it meant no common measure between magnitudes. In this context, magnitude is just another word for value or quantity. Incommensurable units measure concepts that appear to have no common basis (e.g. meters (length), radians (angle), and kilograms (weight) -- all measure different kinds of things, different concepts.

However, incommensurable [measurement] units can have relationships to each other, for instance, in the way that the weight of a substance might be related to its length, but that relationship may not be a simple ratio, as it is with commensurable units.

Insight: A magnet, for example, is a coherent mass with incommensurability of its atomic structure (its lattice work).

Commensurable means "a common measure". It is of course possible for unit names in different systems [of measurement] to measure the same concept. For instance, "feet" and "meters" both measure [the concept] 'length' (linear movement) in a given direction.

With two commensurable units, one unit can be used to measure the magnitude of another unit (e.g., the meter stick can be used to measure the length of the yardstick, both of which represent some specific magnitude of the same concept, length).

> **NOTE:** *Every conversion represents an inefficiency and the possibility for error.*

Commensurable units, because they measure the same concept, can be converted between. The concept, 'conversion [of units]' is the conversion between different units of measurement for the same quantity, typically through the input of a multiplicative quantity known as a 'conversion factor' or 'multiplication factor'.

There are a number of mathematical ways of actually making the conversions, but the one that is most likely to avoid errors involves making a ratio from the conversion units that equals 1.

The method for converting units comes right from one principle:

- Numbers with units (e.g., 16.2 meters or 32 ft/sec²) are treated exactly the same as coefficients with variables (e.g., 16.2x or 32y/z²).
  - Hence, it is not possible to add 32 ft to 32 ft/sec, any more than it is to add 32x to 32x/y. And, when 32 miles is divided (factored as a ratio) by 4 hours to get 8 miles/hour, which is exactly the same (i.e., conveys the same meaning) as dividing 32x by 4y to get 8x/y.

In mathematics, any number can be multiplied by 1, and its value will not change. Multiplying by 1 - a carefully chosen form of 1 - is the principal input required to convert[ing between] units with a different standards of measure[ment], but measuring the same thing (the same concept). A fractional (ratio) form of the real number 1 is required.

For example, imagine the requirement of converting a quantity of hours (e.g., 4 of unit 'hour') to minutes (e.g., ? of unit 'minute'). It is given by the metric system that 60 minutes = 1 hour. When both sides are divided by 1 hour. Herein, the unit hour is treated as a variable. As a variable, 60x = 1y, and both sides can be divided by 1y. After the act of dividing creates a ratio. When, for example, (60 min) / (1 hr) = 1, then any measurement can be multiplied by that fraction and its value does not change. If the quantity of the unit 'hour' is 4, then that quantity (4) is multiplied by the specified ratio form of 1:

- 4hr x (60min / 1hr) =
- (4hr x 60min) / 1hr =
- (4 x 60 min) / 1 =
- 240min
- The initial unit quantity is not a dimensionless pure number (4.0), but is a number with dimensions (4 hours). And, the final result is not a dimensionless pure number (240), but is a number with dimension (240min). The dimension (or measurable concept) is the same for both units. A number with units is different from a number without units or with different units, just as 8x is different from both 8 and 8y. If the top and bottom of the fraction are equal, the fraction equals 1, and the value after multiplying is the same as the value before multiplying—but expressed in different units.

The conversion process has three steps:

1. **Identify conversion equation** - Identify (find and/or determine) a conversion factor between the given units and the desired units, which is expressed as a conversion equation.
   - For instance, 1 mi = 1.61 km or 1 km = 0.621 mi.
2. **Identify conversion ratio/fraction** - Determine the fractional form of the real number 1 by converting that equation to a ratio (fraction) with the desired units on top and the given units on the bottom.
   - For instance, 1.61km/1mi [=1] or 1km/0.621mi [=1]. In this case, the multiplication factor for converting from:
   - mi to km is 1.6 (1.61km/1mi)
   - km to mi is 0.621. (1mi/0.621)
   - Note: If the given units are raised to a power, raise the conversion fraction to that same power.
3. **Multiply** – Calculate the multiplication of the original measurement (the measured quantity as 1 unit of) with the multiplication factor (ratio/fraction), and then, simplify [the units].

In the metric system, the zero point is the same for all units. Some other unit systems set their units zero point to zero too. For instance, 0 pounds equals 0 kilograms, 0 liters equals 0 cubic centimeters, and so on. Take note that between different common unit systems for temperature measurement, is not true: 0 degrees C is a different temperature from 0 degrees F. It is possible to apply the conversion technique to convert between temperature units with different zeros after relating them to a common zero point, and it is more efficient to apply the standard formula as a special case: F = 1.8C + 32. This formula is the slope-intercept form of the equation of a straight line. With other conversions, the intercept is 0 because the conversion line passes through (0,0); but with temperature there's a nonzero intercept because 0 degrees in one measure is not equivalent to 0 degrees in another.

Some conversions are completely impossible, not just impossible using the techniques on this page but impossible by any means at all, because of an axiomatic conceptual contradiction or technical impossibility. For instance, it is not possible to convert 'gallons' to 'square feet' (or liters to square centimeters) using any techniques. This is because gallons and liters measure volume, and square feet or square centimeters measure area. It's like converting x^3 to x^2: it's just not meaningful. A dimensional analysis can be used to show this in a formal way, but informally, remember that area is two dimensions of length and volume is three dimensions of length, and measurements you convert must always have the same number of dimensions.

The following terms mean the same thing: conversion ratio, unit factor, conversion factor, and multiplication factor. This ratio can then be used to multiply the original units to achieve the conversion. Since the ratio = 1 this multiplication does not change the item, it just changes the units.

A conversion ratio (or unit factor) is a ratio [that must be] equal to one. This ratio carries the names of the units to be used in the conversion.

1. **Factor** - It is a determining factor in the conversion.

2. **Ratio** - It is a ratio that carries the names of the units to be used in the conversion.
3. **Unit** - All conversion ratios (unit factors) must equal one.
4. **Multiplication** – The unit quantities are multiplied -- input of a multiplicative quantity. Multiply the measurement (# units you have) by the conversion ratio.

A conversion factor is a ratio (or fraction) that represents the relationship between two different units. A conversion factor or multiplication factor, originally known as 'unity bracket method', is a mathematical tool (a method) for converting between different units of measurement. It is sometimes referred to as a 'unit multiplier'. The method involves a ratio (fraction) in which the denominator is equal to the numerator. The conversion ratio is based upon the concept of 'equivalent values'.

A conversion factor is [a quantity] used to change the units of a measured quantity without changing its value (i.e., its known quantity). Because of the 'identity property' of multiplication, the value of a number will not change as long as it is multiplied by one. Also, if the numerator and denominator of a ratio (fraction) are equal to each other, then the ratio (fraction) is equal to one. So as long as the numerator and denominator of the ratio (fraction) are equivalent, they will not affect the value of the measured quantity.

For example, the unit [of measurement] 'days' may be converted to the unit [of measurement] 'hours', by multiplying the 'days' by the conversion factor 24 (a quantity).

Conversion factor examples include:

- Quantity = [set equal to]= 1 day = 24 hours = 1440 minutes; therefore, 15 minutes (1 day/1440 minutes) = 15/1440 ~= 0.010416667 = ~0.01 days.
- Quantity = [set equal to]= 1 hour = 60 mins = 3600 seconds; therefore, 7200 seconds = 120 mins = 2 hours.

Some unit systems do not have a common basis for their conversion/multiplication factor. In the metric system, however, conversion between units can be discerned by their prefixes (for example, 1 kilogram = 1000 grams, 1 milligram = 0.001 grams). Precision of language is important, and the presence of exceptions (e.g., 1 micron = 10−6 metre) are likely to cause confusion.

4.13 Instrumentation

NOTE: *Measurement instruments are devices that replace the need for actual measuring units (i.e., objects) in making comparisons.*

There are three measurement instrumentation system unit types:

1. **Device** - device measurement units are the units of measure of the output device.
2. **Absolute** - absolute measurement units are units that do not depend on the device. For example: inches and meters are absolute units, because their length does not depend on the output device.
3. **Relative** - relative measurement units depend on the size of "something else". For instance, a measurement system's units may depend on the size of the parent or root objects containing the measurable object.

5 Numbers

APHORISM: *All is number.*

Number is a syntactic category. A syntactic category is a type of syntactic unit that theories of syntax assume. Word classes, largely corresponding to traditional parts of speech (e.g. noun, verb, preposition, etc.), are syntactic categories. In phrase structure grammars, the phrasal categories (e.g. noun phrase, verb phrase, prepositional phrase, etc.) are also syntactic categories. Dependency grammars, however, do not acknowledge phrasal categories (at least not in the traditional sense).

The word 'number' belongs to a noun of multitude standing either for a single entity or for the individuals making the whole. An amount in general is expressed by a special class of words called identifiers, indefinite and definite and quantifiers, definite and indefinite. The amount may be expressed by: singular form and plural from, ordinal numbers before a count noun singular (first, second, third...), the demonstratives; definite and indefinite numbers and measurements (hundred/hundreds, million/millions), or cardinal numbers before count nouns. The set of language quantifiers covers "a few, a great number, many, several (for count names); a bit of, a little, less, a great deal (amount) of, much (for mass names); all, plenty of, a lot of, enough, more, most, some, any, both, each, either, neither, every, no". For the complex case of unidentified amounts, the parts and examples of a mass are indicated with respect to the following: a measure of a mass (two kilos of rice and twenty bottles of milk or ten pieces of paper); a piece or part of a mass (part, element, atom, item, article, drop); or a shape of a container (a basket, box, case, cup, bottle, vessel, jar).

5.1 Number connections

Notational connections, and the representation of relationships within the concept 'number'.

In mathematics, the concept of 'number' has the following expressed relationships/attributes:

1. **Conception** – natural (counting), whole (direct), integer (inverse), ir/rational, real, imaginary (lateral), complex (angular).
2. **Notation** – the methodological expression of a number.
   A. **Symbol representation** – a sign of operation and/or a representation of a constant [number].
   B. **Numeral representation** – the digits of the numeral system.
   C. **Radix** – the cardinal of the [non-repeated] sequence of digits in the number system.

Every number has the following two attributes:

1. **Value** - The number a numeral represents is called its value.
2. **Sign** - The dimensional direction of the value. In general, there are two signs, positive (+) and negative (-). The number zero (0) has no sign and may be considered to have a neutral sign.

5.2 Initial mathematical relationships

When there are two numbers in awareness/memory, the larger will always lie to the right or left (up or down) of the smaller one. Symbols are used to communicate relationships between numbers [on the number line].

The following relationships allow for integrated comparisons between numbers.

There are equality relationships:

- = is equal to, of equal value
- ≠ does not equal, is not of equal value
- ≈ is approximately equal to
- in Mathematica: Set (=) vs Equal (==) vs (identically) SameQ (===).

There are order relationships:

- < less than
- > greater than
- ≤ less than or equal to
- ≥ greater than or equal to

In logic, and hence, mathematics, there is an order to operations (i.e., an order to relationships). In mathematics, this order is known as 'the order of operations'.

The operation minus from the sign of a negative number (as -2 = negative 2).

5.3 Number

A number is a count or measurement, that is really an idea in our minds, which may represent a state or condition of the real world.

We write or talk about numbers using numerals such as "4" or "four". We could also hold up 4 fingers, or tap the ground 4 times. These are all different ways of referring to the same number.

INSIGHT: *Number represents movement.*

There are also special numbers (like π (Pi)) that can't be written exactly, but are still numbers because we know the idea behind them.

A number is [the value or count] of a set of something similar. It is the conceptual expression of an iterating pattern. A given pattern may or may not exist, and if it exists, then how many iterations of that pattern exist. Geometrically speaking, it may also be said that a number is the “sum” of identical (indistinguishable) fractal points. The concept “sum” introduces a mathematical concept/unit, sum (or algebraic total of that which is indistinguishable).

As a concept, ‘number’ represents the presence of the iteration of information (i.e., the presence of patter). Once a pattern is present, logic (as math[ematics]) can be applied to process (calculate). A number is what satisfies the axioms of its number system. In mathematics, a number is a mathematical object used to count, measure, and label.

Numbering and mathematical logic are used to model and understand the universe:

1. To number is to understand iteration.
2. To map is to understand relationship.
3. To calculate is to understand creation.
4. To articulate is to create.

> **INSIGHT:** *The primary function of numbering object and spaces in community is for identification and wayfinding. Numbering allows for coherent creation and dis-creation.*

When a sensation is being measured with numbers, the understanding is language independent, and numeration is language dependent visualized as a specific linguistic expression.

The finite number of digits used in the numbering process is called the **radix/base** [of the selected number system].

Each characterized conceptualization of a number involves the linguistic/logical creation of a ‘mathematical construct[ive]’ operation (or process). Some of the following characterizations represent groups of constructs. Within that which is termed the “real” number system, there is an increasing order of mathematical constructive complexity, moving from natural numbers at the lowest order, to rational/irrational at the highest order. The misnamed, “imaginary” numbers, represent the extension of the number system into a second, angular (perpendicular) dimension.

Some of the conceptualizations can be viewed as levels or dimensions. For 1 dimension, 1 piece of information is required to define location in that dimension. A real number is sufficient to define location in one dimension. For 2 dimensions, 2 pieces of information are required. Either 2 real numbers or a complex number (note: a complex number holds two pieces of information in a single “number”). If only real numbers are used, then a vector matrix must be generated to hold multiple pieces of information. A vector matrix is not necessary for complex numbers. For both the real vectors and the complex numbers, operations must be defined for translation, rotation, and integration, and duplication (or subtraction). For 3 dimensions, it is possible to use a 1X3 matrix with 3 real numbers in it, or extend the complex numbers to something more complex. The complex number must hold 3 pieces of information, rather than 2. The translation and rotation operations are re-encoded (redefined) to work within each system. To move from 2d to 3d, a new operation is created with another “imaginary” axis. For example of a 2d complex number would be 2+3i. A 3d complex number will be 2+3i+4j. A 4 d number would be 2+3i+4j+5k. The math becomes a lot more complicated ass you add dimensions, and you need computers to do it. Plus humans cannot visualize in more than 3 dimensions. However, the core math stays the same. The imaginary notations are merely notations. Group theory creates different number systems, from natural numbers N, to integers Z, to rationals Q, to reals R, and complex plane C, and on to higher dimensions.

It should be possible to agree on this (*):

- N ⊂ Z ⊂ Q ⊂ R,
  - where, ⊂ means subset
  - When A is a subset of B (A ⊂ B or A subset B), then saying that x is in A implies that it is in B, but not necessarily conversely. Who focuses on A, and forgets about B, may protest against a person who discusses B.
  - When it is said that the rational numbers are “numbers”, because they are in R, then group theorists might protest -- the rationals are “only” numbers because: (1) Q is an extension of Z, by including [the mathematical operation of] division; and (2) “we” call them “number” too.
- The elements in R are called numbers.
- The elements in the subsets are called numbers too.

A set (or system) is a collection of objects, typically grouped within braces { }, where each object is called an element (part or sub-system). For example, {red, green, blue} is a set of colors. A subset is a set consisting of elements that belong to a given set. For example, {green, blue} is a subset of the color set above. A set with no elements is called the empty set and has its own special notation, { } or Ø.

5.4 A number system

The real number system represents fully functional field for mapping the real world. A number system is a set of objects (often numbers), operations, and the rules governing those operations.

- A unary [number] system has one numeral in the set {1}.
- A binary [number] system has two numerals in the set {1,2}.
- A trinary [number] system has three numerals in the set {1,2,3}.

- The decimal [number] system has ten numerals in the set {0,1,2,3,4,5,6,7,8,9}.

A number system is the logical composition of sets of symbolic [numeral] digits, that are used to represent the possible enumerations of the concept of 'number'. The system is the concept 'number', which is decomposed into mathematically operative, numerical subsets [of the unified 'number' system set].

The "natural" [counting] numbers form the first set of numbers upon which [mathematical] operations can be performed. This number system begins with the conception of natural counting numbers.

> **AXIOM:** *Counting numbers are the origin of all numbers. Ordered pairs exist – a sequence of patterns exists. ∴ Complex numbers exist. ∴ Complex thoughts exist.*

The concept of 'number' is composed of [at least] the following characterizations, which are otherwise, conceptualizations of the concept, 'number'. Together, the possible numerals at each level of conceptualization are known as a 'set' (mathematical). The sets of numbers in the "real world" number system include, but may not be limited to:

1. **[N] A natural [counting] number (positive integers)** – the concept of, [the numerical mapping of] pattern recognition as a finite sequence of digits, through which an infinite sequence of numbers (numerical meaning) may exist. A finite sequence of digits representing an order of iteration. For example, 1,2,3,4,5,6,7,8,9. There are infinitely many natural numbers; the set of natural numbers is infinite -- as is the set of all squared numbers, the even numbers, the odd numbers, the rational numbers, and the irrational numbers. There is a hierarchy of these infinities, the so-called transfinite numbers. In counting, someone can simply keep adding 1 to the previous number to get more and more. Natural numbers are investigated in an area of mathematics called Number theory.
   - Counting numbers are actual symbols that can be visually expressed and used to represent numbers. Counting numbers are now called positive whole numbers.
   - Number has 'value'.
   - The set of natural numbers.

2. **[W] A whole [counting] number (zero integer)** – the concept of the absence of that which is being counted, together with the principle of a natural number. The absence is commonly expressed with (represented by) the symbol (digit), 0. There is one addition to the finite number of digits in the natural counting conceptualization of 'number', the absence of the pattern being sequenced. Depending on perspective, zero may be considered unsigned, or may be considered its own sign. Note that the first such "unreal" in the "real" versus "unreal" paradigm was the zero.
   - Number has 'magnitude'.
   - The set of whole numbers.

3. **[Z] An integers number (negative integer)** – the concept of a opposite (i.e., different state, direction, or reverse) applied to the sequencing pattern, together with the principles of a natural counting number and a whole number. Sign of a direction - it is common to label certain directions as positive (+, or nothing) or negative (-). Negative numbers are the next most obvious addition, as [in part] the representation of a reverse in direction. If something is not a whole number, then it is not an integer. In other words, an integer can be negative, positive, or zero; and it is at the integer level of the characterized conceptualization of number that the concept of a 'negative integer' is added. The conceptualization level is 'integer', but the new conceptualization at this level is the 'negative integer'. A negative integer has a sign in front of the digit, -4,-3,-2,-1. Besides zero (when it is considered a sign), the concept of sign originates from the property of there being a possible "polar" difference in any given number representing the presence of the pattern. The idea of a "change of sign/state" is used throughout mathematics and physics to denote the additive inverse (negation, or multiplication by −1). Note that a decimal number (e.g., 132.493) is not an integer. Integers can be added, subtracted, and multiplied. In application, the negative symbol is just a relative symbol, not an absolute value. In other words, "-5" is not conceived as "negative 5", but as something opposite of something else.
   - What does it mean to subtract (take away) a larger positive integer from a smaller.
   - Negative integers double the number of elements/digits present.
   - Number has 'sign'.
   - The set of integer numbers. There are infinitely many integers sequencing in two opposite directions.
   - With the conception of "negative" (inverse) and "positive"(direct), the concept of "opposite" arises (e.g., the opposite of -3 is 3; or, the opposite of unsigned numeral A, is signed numeral A). The "double-negative" property says that the opposite of a negative number is not a negative number (i.e., the opposite of -7 is –(-7).

4. **[R] A real number (a field, plane)** – is a value that represents the quantity [of a sequence] along a single dimension (a line). In between any two given real numbers there exists an infinity of real numbers. Real numbers can be visualized as points on an infinitely long number line. The word “real” was historically introduced to distinguish between the real and imaginary roots of polynomials. The term, ‘polynomial’ comes from poly- ‘many,’ on the pattern of multinomial (a pattern named “term”). A polynomial is an expression of more than two algebraic terms. In other words, a polynomial is an expression consisting of variables (or indeterminates) and coefficients, that involves only the operations of addition, subtraction, multiplication, and non-negative integer exponents. Polynomials are used to form polynomial equations, which encode a wide range of problems.
    - In mathematics, a plane is a flat, two-dimensional surface that extends infinitely far. A plane is the two-dimensional analogue of a point (zero dimensions), a line (one dimension) and three-dimensional space.
    - A field is a two-dimensional plane with the natural addition of vectors.
    - The real number line/set. The mathematical terms line and set, as applied to the real numbers, come from two different philosophical approaches to knowing and naming things. The term line as a representation of the real numbers, such as in the real number line, descended from geometry (Euclid), whereas the use of the term set as a representation of the real numbers, descended from algebra, and specifically set theory, introduced by Cantor.
    - A line has no thickness, because a line is a mathematical construct conceived as a tightly strung string of points (data packets) formed by the junction of two planes, where an infinite subdivision could occur.

5. **[Q] A rational [ratio-nal] number (fractions of prior numbers)** – a ratio (fraction) of any of the individual prior conceptualization, with the exception that 0 is not ever a denominator. Rational numbers are quotients (Q), which are the result of division (i.e., sub-division). In a fraction, the denominator represents the number of equal parts in a whole, and the numerator represents how many parts are signified (i.e., “being considered”). In other words, a ratio of a positive (e.g., using the 1 integer: 1/1, -1/1, 0/1 (is 1), but never 1/0 (which has no meaning and an undefined result). The number above and below are members of the same integer set. A decimal number (1.5) is a rational number because the digit(s) to the right of the symbol are just another way of writing a ratio (fraction). Rational numbers are integers, and fractions of integers, put together, but there are not any other numbers. All rational numbers may be represented in radix point or fractional form. Fractions/ratios have three forms of notation: radix point notation; fractional/ratio notation; and graph (visual) notation.
    A. **Radix point (e.g., decimal) notation** – for example, 0.5. The decimal expansion of rational numbers is either finite (like 0.73), or it eventually consists of repeating blocks of digits (like 0.73454545...).
    B. **Fractional/ratio notation** – for example, 1/2 or 1:2.
        1. The expression 1 / 2 represents both the operation of division and the resulting number. This is an example of a “procept“, the combination of process and concept (More completely, a procept is an amalgam of three components: a process which produces a mathematical object and a symbol which is used to represent either process or object.).
        2. The procept property of y/x . In common parlance, there are separate definitional entries for division, quotient, fraction, ration, and proportionality. This is an inconsistent nomenclature. Instead, ratio is the input of division, and number is the result of division. This is not a definition of number but a distinction between input and output of division. It is suggested to use the terms, (static) quotient, for the form with numerator y “divided by” denominator x.
        3. Multiplication is not a precept; there is a clear distinction between the operation 2 • 3, and the resulting number 6. It is also logical to say that 2 • 3 = 3 • 2. The word “multiplication” could be completely replaced by the term ‘group’, and ‘grouping’. With 6 identical elements, there can be organized 3 groups of 2.
        4. (static) quotient[y, x] = y / x
    C. **Graphing** – this content is well visualized as a graph. Here, it could be said that there is the notion of proportional space: {denominator x, numerator y}
    D. In a two number line graph, the denominator (cause) on the horizontal axis and the numerator (effect) on the vertical axis (instead of reversed), as it should be because of the difference quotient in calculus.

Number as an internally iterating sub-pattern, which may be representable either by a finite number of digits or a [forever] repeating pattern of digits. Number has sub-division.

The language becomes confusing here because, to say, rational seems to indicate that the fractions were somehow qualitatively better, or at least more rational than the irrational numbers

The set of rational numbers -- rationals/fractions are countably infinite. It is often suggested to abolish the word "fraction". Number means also satisfying a standard form. Thus "number" is not something mysterious but is a form, like the other forms, yet standardised. For example, we have 2 / 4 = 1 / 2, yet 1 / 2 has the standard form of the rationals so that 2 / 4 needs to be simplified by eliminating common prime factors. The standard form (prime factor) of 2/4 is ½. For a standard form for the rationals, the rules are targeted at facilitating the location on the number line, while we distinguish the operation minus from the sign of a negative number (as -2 = negative 2).

If a rational number is equal to an integer, it is written as this integer, and otherwise:

1. The rational number is written as an integer plus or minus a quotient of natural numbers.
2. The integer part is not written when it is 0, unless the quotient part is 0 too (and then the whole is the integer 0).
3. The quotient part has a denominator that isn't 0 or 1.
4. The quotient part is not written when the numerator is 0 (and then the whole is an integer).
5. The quotient part consists of a quotient (form) with an (absolute) value smaller than 1.
6. The quotient part is simplified by elimination of common primes.
7. When the integer part is 0 then plus is not written and minus is transformed into the negative sign written before the quotient part.
8. When the integer part is nonzero then there is plus or minus for the quotient part in the same direction as the sign of the integer part (reasoning in the same direction).

6. **An irrational [ir-ratio-nal] number (an irrational root)** – a [real] number that cannot be expressed as a fraction. Irrational, but can be expressed on a number line. Some numbers cannot be written as a ratio of two integers; they cannot be expressed as a fraction of integers (non-fractions). It was discovered that the square root of 2 cannot be written as a fraction. Neither π (pi) nor e can be written as fractions. Note that Cantor, the inventor of set theory, published a paper defining irrational numbers as convergent sequences of rational numbers. What is being observed in the category called the "irrational" (i.e., the irrational numbers) are the numbers that "fill in" all the "gaps" between the rationals [on a number line]. Irrational numbers are those which can't be written as a fraction (which don't have a repeating decimal expansion). Those rational numbers which aren't the result of polynomial equations with rational coefficients.
   A. The diagonal of a unit square cannot be represented by a ratio of two integers. And yet, this number does have a direct geometric representation [in the number system].
   B. Rationals are the separation of the whole into parts, and irrationals are a unique category of this type of separation.
   C. Irrationals are uncountably infinite.
   D. The set of irrational numbers.

7. **[C] A complex number (a complex field)** – a field that extends the "real" field, 2D visualization) - A set of real numbers and non-real numbers put together would be a "complex" set of numbers, and hence, the concept given to this category is 'complex'. Complex numbers answer the problem of determining the square root of negative real numbers. Complex numbers are the final step in a sequence of increasingly "unreal" extensions to the [natural] number system that humans have found it necessary to add over the centuries in order to express increasingly sub-divided, significant, numerical concepts. Complex numbers are typically represented graphically as points in the 2D plane, and the rules of addition and multiplication are equivalent to certain operations on lengths and angles. A complex number is a number that has both a Real and an Imaginary part. That is it has 'length' residing along the Real number line (the usual numbers we're all familiar with), what we call the Real axis, and a 'height' residing along an axis perpendicular to that Real number line, which we call the Imaginary axis. In mathematics, the complex plane or z-plane is a geometric representation of the complex numbers established by the "real" axis and the perpendicular "imaginary" axis. This is visualized as a modified Cartesian plane, with the real part of a complex number represented by a displacement along the x-axis, and the imaginary part by a displacement along the y-axis. The concept of the complex plane

allows a geometric interpretation of complex numbers. Here, it is the "function" that [visually] distributes the number along the z-axis. Complex numbers form a two-dimensional "vector space" over the "real" numbers. Therein, for each complex number c, there exists real numbers a and b such that c=a + i*b. Here, complex numbers are visualized as a two dimensional plane (the "regular" axis, and the "I" axis). Complex numbers have the added property that rotation in this "plane" is simply a multiplication by a complex number (this, too, has a rigid mathematical definition, that "every R-automorphism is represented by multiplication". Let's call these 3 properties 1,2 and 3.

- Just like coordinates can be plotted on the x,y plane, complex numbers are represented in the complex plane. In a normal x,y plane there is no connection between the two dimensions; there are there are no rules about how they can relate to one another. In a complex plane, there are the rules of algebra. i has to do with rotation on a complex plane. Angles can be determined through the additional use of trigonometry.
- The i as a notation of a complex number, and using notation makes the math more simple.
- It is a field.
- It is two-dimensional vector-space over the reals.
- At every rotation in the vector space is represented by multiplication.
- Every non-contradictory equation, algebraic or transcendental, has a solution within the application of complex numbers. The present understanding is that their addition finalizes (or wholes) the number system into a self-sufficient, consistent system.
- Complex numbers exist in the same way "real" number exists, they're involved in our daily technically computational societal operations. Complex numbers are used for the representation of various physical phenomena, including states of particles and the behavior of electrical currents. They are also necessary (or, at least, efficient) at computing 3D visual space (i.e., computer graphics). All electric and magnetic systems behave like complex numbers express in numerical and graphical notation.
- The Cartesian coordinate system but two such real number lines drawn orthogonally across one another at what is called the origin, used to visualize the mapping of a real function from one real number line called x to another "real number line" called y. The rest of mathematics is supported upon this foundation [of operation].
- Complex numbers are best represented in a coordinate system were the x-axis shows the real part and the y-axis shows the imaginary part of the complex number. Therein, complex numbers use the Cartesian (x, y; x +iy) coordinates, or use an angle and the distance from a fixed point (the origin) as polar coordinates (rei0).
- The polar equation becomes 1 x e π × i = -1, or e π × i + 1 = 0. This equation is significant because it involves all the fundamental constants in mathematics: 0,1,e,π, and i. These numbers are a numerical-conceptual mapping to relationships in the existent world.
- A complex number is a point on the two dimensional "field" plane of a real number. The absolute value of a complex number is its distance to the origin, and let's call its "angle" the angle it forms with the positive x-axis. The Real numbers are just the x-axis, and "i" is just (0,1). So the real number 1 is (1,0) and -1 is (-1,0). Then multiplying complex numbers multiplies their absolute values and adds the angles. That's why (0,1) times itself is -1 = (-1,0). 90 degrees plus 90 degrees = 180 degrees.

INSIGHT: *In mathematics, the complex numbers are the final step/level/order in the sequenced system of counting numbers [as patterned iterations). From this perspective, every other number is just a sub-set of the complex numbers; the complex numbers are that which is presently understood to be real, and an existent part of the functionally real world.*

8. **A transcendental number** – a transcendental number is a [possibly] complex number that is not an algebraic number—that is, not a root (i.e., solution) of a nonzero polynomial equation with integer coefficients. Hence, a transcendental number is not a finite or repeating set of digits, and also not representable as a root. It is real or complex number that is not algebraic – that is, it is not a root of a non-zero polynomial equation with integer (or, equivalent, ration) coefficients. Transcendental numbers can't be defined as the solution to a specific algebraic equation. Every real transcendental number must also be irrational, since a rational number is, by definition, an algebraic number of degree one.
    - For example, π (pi) - 3.1415926535897...
    - For example, e (Euler's number) is the base of the natural logarithm - the unique number whose natural logarithm is equal to one - 2.71828.
    - Everything that exists follows eternal rules describable as ratios of numbers. Thus, any number could be written as a ratio. For example, 5 as 5/1 or 0.5 as 1/2. Even a number with an

infinite decimal sequence can exist as a ratio. All of these are rational numbers. Historically, one number was found to violate this rule. A square with each side measuring 1 unit according to Pythagoras theorem: a2 + b2 = c2. The diagonal of the square length would be square root (sqrt) of 2. Square root of 2 cannot be expressed as a ratio of two integers (e.g., 45/34 or 33/283), and is thus an irrational number because it can't be written as a ratio of two integers.

9. **[i] An imaginary number** - Imaginary numbers are conceived of as numbers, but are not real numbers. Literally, imagined numbers. “Imaginary” numbers visualize a 3d graph. The poorly named “imaginary” numbers (number system) represent the second dimensional (or angular) data point (coordinate) of a number – here, formal relationships can now be constructed (and functions generated) between the two axes (dimensions). All of the [other] “real” numbers are presentable on a one dimensional number line. “Imaginary” numbers exist to extend the functional [mathematical] operation of the numbering system. The poorly named “imaginary” numbers represent an angular separation from the one dimensional number line into a two dimensional numbering system, with the second dimension existing, and capable of being visualized, as perpendicular to the first. Because it couldn’t initially be conceived of possibly existing in the real number world it was given the name imaginary or impossible. In response to this understanding, the other set of numbers gets called real. And, when a number contains these two parts, it is called a complex number. Imaginary numbers are not on the “real” number line, but in mathematics, they are just as real as any other formally conceptualized number. Where negatives rotate in 180 degrees , i rotates 90 degrees. i^2 is -1.
    - Imaginary do not exist apart from the real numbers, but exist in what is conceived of as a 90°perpendicular dimension. They are the natural extension of the number system from one to two dimensions. Numbers can be conceived of as two dimensional.
    - The imaginary number [set] – there is only one number here.

Unique conceptual classes of numbers include, but are not limited to:

- Natural number level - Even and odd – every other sequence, with even and odd, existing one sequence off of the other.
- Ratio level - Prime numbers – are numbers whose factors are 1 and that number. A number that has more than three numbers that go into that number is the opposite. 7 and 3 are prime.
- Natural number level - Infinite numbers – the smallest of which is the number of integers, represented by ℵ0.
- Division-algebra (quaternions octonions)
- Going to 4D (i.e. Quaternion), multiplication is no longer commutative. Going to 8D (i.e., octonions).
- Sign of angle, the “degree” of change between two subdivisions.
- Sign of change - the delta symbol, Δ.

5.5 *The zero as unity perspective*

> **NOTE:** *The cycling process is called iteration. The output of on operation becomes the input of another, and so on.*

The path the earth follows as it spirals around the sun can be described as 0, a circle, and however, the year is divided, each division is part of the whole 0 cycle. The earth itself rotates around an axis, a circular motion that defines a day as the passage of the risen sun to sunset and back to sunrise.

The number of days can be described as 1 each, or as sets of days, as weeks or months, and in a tangible and quite natural way, the circle, cycle of 0, can be said to be the natural phenomenon that allows each 1 day to exist, just as the earth’s cycle around the sun allows a year to exist.

Without the 0 cyclical movement 1 doesn’t exist and 1 cycle isn’t complete until the circle 0 has been drawn. In both these natural instances, 0 can be said to be a unified 1 or the completion of a cyclical movement that is now designated as 1 symbolic and naturally significant 0. One (1) earth completes each of the circles that describe a visual 0 path, and each 0 has a tangible form that is described by 1 earth and 1 sun. So, does 1 + 1 = 0 or does 0 = 1 + 1.

Does the cycle 0 exist before that which is describing it, or does the cycle 0 only exist because 1 + 1 brings it to the attention of human minds? Zero (0) is a symbol, a tangible cyclical passage that describes unity, but it also describes the completion of 1 day or 1 year.

The end of a year does not result in nothing any more than the end of a day does, they both bring a new cycle, a new 0, that will be unified until the current natural course of events principally determines that the cycle is no longer sustainable, and for the purposes of community, fulfilling. Through subsequent divisions, new 0s, cycles will be formed (as 10,100,1000,10000). In math, these cycles are known as “orders of” [the concept] ‘magnitude’. Each cycle is another “level” (a.k.a., “order”) of the iteration (as in, a whole new sub-division).

The minds that first described a 0, what we now call zero, didn’t imagine it, 0 was real as a unifying circle or cycle, and the concept of someone taking all the candies to leave nothing described something else.

The process of subdividing a fixed line (concept) infinitely leads to something that is infinitely long - the Mandelbrot set. More ("finer") divisions are always possible, and the "finer" the view, the greater the number of divisions. However, mathematically, something which is infinite cannot be measured. Hence, everything which is to be measured must be conceptually and/or mathematically bound.

Non-Euclidean geometry is a consistent system of definitions, assumptions, and proofs that describe objects as points, lines, and planes.

A fractal is something that is self-similar. It is an operation (mathematical process) involving the iteration of an equation with an input coming from an output. Take a number, process it through the 'formula', the result is a number, which is then fed back into the formula. And, that whole process is consecutively iterated, over and over again. What happens when this occurs many times? The structure of numbers that appears is called a 'set' (the Julius set in particular). The Julius set can be visualized (as a whole) on a graph with two axes representing two scales with at least 1 fixed reference point (i.e., with at least one self-similar subdivision). This point represents their unification.

It is possible to understand this, because in mathematics there is a system of definitions, assumptions, and proofs that describe objects as points, lines, and planes (a.k.a., non-Euclidean geometry).

The perspective can be on a one-to-one basis or a one-less-one basis. In concern to a one-less-one basis, the first condition has no condition before it, and hence, zero is conceived. The second one (condition) has one (condition) before it. The third one (condition) has two (conditions) before it. And so on.

The decimal numeral system is a column system until 9 units transcend into zero, a new column. A unifying zero is used because what has gone before is agreed on, and the next column starts with the transcended one, as one ten.

5.6 Continuum

A continuum is a whole of differentiated parts, where the differentiated parts have continuity. The parts are intrinsically differentiated by their relation to the whole, which forms a unified whole. (and hence, the whole is whole/unified). The various analogical uses of the term continuity express the extensiveness of the concept. Relative to the primary analogate, which is the extensive continuum, the parts may be understood under two formalities (eventually becoming formulas, expressions, and arguments):

1. The analytic [part]: a whole that is divisible without end into (analytic) parts, of which there is no smallest. This is interval (the pattern sensation) part.
2. The compositive [part]: a whole, the extremities of whose (compositive) parts are one. This is the sequence part, the iteration. This is a composite of [the following information inquiry processes], which can return data or no data:
   - How many sequential iterations are possible?
   - How many sequential iterations are associated ("take up by") the pattern?
   - Where in the possible sequence of iterating patterns do the intervals start (initiate) and stop (terminate)?

Continuity (without separation, continuous in existence) is logically related to contiguity (bounded in existence, direct contact, complete separation, axiom) and consecutiveness (sequence of separation/occurrence, order); all three (continuity, contiguity, and consecutiveness) refer to extension, but constitute a different ordering of parts. There is the sensation of a whole, and then, its parts and the sequential ordering of those parts.

A whole sub-divides (i.e., "has separation") if some element of a different [extensive] nature integrates ("intervenes") between any part of the ordered whole, to form an interval of parts. Note that it is presently thought that continua (multiple continuous extents, series, or wholes) are encoded into humans cognition means of a scanning motion [of the extensional sensation sub-system].

In other words, a continuum is essentially one, though it has distinguishable parts, whereas both contiguous and consecutive entities are pluralities only, the former with parts distinguished and bounded (separation), the latter with parts separated in sequence (order).

The origin of the notion of continuum is most readily traced to the sensible experience of physical extension, which is the first formal effect of dimensional quantification, manifesting factual material unity but remaining subject to division. "I am that I am" is a definitional phrase for "pattern and sequence", which may be numerically notated as "1,2,3,1,2".

All continua are divisible into parts, that are themselves divisible continua; and since such division does not add anything, the positions of the divisions must be marked by pre-contained indivisibles. Division is generally not considered as a form of separation. The actual division of an abstract continuum is accomplished by the mental removal of a portion of the continuum or the establishment of an indivisible boundary, which becomes a categorically designated separation. Division produces two continua that are either contiguous or consecutive. Boundaries on a continuum of composite parts are resolved by situ or position (the differentia of dimensive quantity). When there are positions (more than one position), then there is geometry in a pattern (or at least, the possible expression of a geometric pattern).

Here, each pattern represents an extensive magnitude, a 'value'. That 'value' position may be fixed as local [motion; i.e., position], or provide relative motion and/or time as a continuum (i.e., relative position).

Though attained in sensation, the abstracted concept of continuum is primarily mathematical; when used in a physical sense, the term is a secondary analogate. It is possible to unity notational priority in the mathematical continuum with experiential priority in the physical continuum. Visually, abstractions are first represented by lines, then angled lines, then surfaces, then three dimensional surfaces, then the iteration of dimensional surfaces (as time).

> **INSIGHT:** *Disagreements about physical existence are often traced to the problem of experiencing reality without an interruption of adaptive continuity.*

There are two types of continuum from the perspective of consciousness:

- In a static continuum, all parts coexist and are known immediately
- In a flowing continuum, the parts, successive in existence, are known only through the re-presentations of memory.

Both types are understood as wholes divisible into parts, distinguished but not interrupted, which have the same nature as the whole; but a flowing continuum is a becoming, hence its parts are never a being, even when considered abstractly.

The mathematical representation of flowing continua requires greater abstraction than that of static continua, since mathematics abstracts from motion. The mathematical notion is subject to further analogical extensions within that order.

Mathematics has one antiquated original separation forming arithmetic (the science of the discrete) and geometry (the science of the continuous). Today, analytical geometry is their unification.

A reference forms a one-to-one corresponding convention to a pattern on an existing continuum. A sensed referential similarity forms a numerical 1 to 1 relationship mapping. From the instantiation of this initial relationship map comes [models and operations for] counting, fractioning, and then, [sub-]dimensioning.

5.6.1 Counting principles

Numerosity is the ability to discriminate arrays of objects on the basis of the quantity of items present. For example, being aware that a quantity of two is different than a quantity of three [of some object/thing].

Five principles defined counting as the process of identifying and operating (functioning) with patterns.

Remember the cooperation from rules theorem:

1. Ideas are [becoming] concepts.
2. Concepts are divided patterns (objects and relationships).
3. Patterns are structured through principles.
4. Principles are "rules".
5. "Rules" are state[ment]s of action.
6. Action is motion.
7. Motion is the result of the instantiation of a resolved information set (space), a decision.
8. An information set (space) resolves into a decision through the logical processing of present information. There may be internal (sub-system, subjective) and external (system, objective) logic here.
9. Logic is pattern recognition.
10. Pattern recognition is awareness.
11. Awareness is the totality of "your" present experience.
12. Counting (sensation) and sequencing (creation) are the basis of technical understanding.
13. Pattern recognition is the basis of intelligence.
14. Logic is the basis of conceptual understanding.
15. Decisioning is the basis of optimization.
16. Motion is the basis of creation.
17. Action is the basis of sensation.
18. Rules are the basis coordination.
19. Principles are the basis of cooperation.

Simplistically, 'abstraction' is the act of giving a short and easy to remember name to something that is long and complicated. By doing this, you absolve yourself of needing to remember the long and complicated stuff. "Abstraction" is the one of the bases of computer science and information processing. In computation, bit patterns represent instructions (operations, processes) for [at least]:

1. Load [an iteration]
2. Store [an iteration]
3. Add [an iteration]
4. Multiply [an iteration]

Bit patterns are hard for humans to remember; hence, they are further encode as assembly language mnemonics (note: assembly language is a base-16, sexadecimal/hexadecimal, number system). Operating systems further abstract the physical hardware that might be connected to a computer, in order to extend functionality (i.e., make it more easily shareable by multiple applications).

For example, in computation, virtual memory is commonly thought of as "paging", which it is, but there is more to the conceptualization. All physical hardware defines a fixed set of categories (names) where program data can be stored. These names/categories are the "physical addresses". If application-programs were forced to always use physical addresses they would constantly have to interrupt/disrupt each other, from which errors are the result.

Virtual memory allows every program to virtually work with some fixed set of addresses, that start at zero, and increase in a finite or infinite . Further, they can pretend that they have this whole address space to themselves. The operating system then takes care of making sure that programs don't end up using the same physical memory or otherwise destroying each other. Again, we simplify high level code by introducing an abstraction layer that essentially does name translation.

The higher-level software applications also have their own collections of abstractions. Some are important enough to have unique names. Most of these applications are constructed out of three axiomatic abstractions:

1. **Model** – how the application stores the basic data (units).
2. **View** – how the application displays the data for the user (entity higher up the supra-systems decision resolution hierarchy). This is the visible part of the user interface (or, the information that is accessible).
3. **Controller** – how the application responds to commands. This is the less visible part of the user interface. It determines which sequences of actions are possible, and thus, what workflows it can support.

In an non-unified model, abstraction boundaries tend to leak out, which is what makes fixing systems in a non-unified model complicated.

An **array** is a systematic arrangement of similar objects – a data structure that contains groups of elements (information sets). When an array is composed of numbers, then those numbers are usually presented in a row and column, matrix, [notational] format. An array is commonly signed with, { } (although there are a variety of other signs).

With the above understandings in mind, the five principles (or principle conceptualizations) of counting are:

1. **The one-to-one principle (a.k.a., one-one principle)** – Assign a single tag (reference association, label, name, category, value, number, word, sign, symbol) to each counted or sequenced object (item, thing, event) in the array [of categorically similar patterns]. In other words, assign only one name/label to each individually counted/sequenced pattern. The two processes required here, to be performed on the collection of objects, are partitioning and tagging. Every item being counted needs to be transferred from the to-be-counted category to the counted category (partitioning), while a distinct tag must be logically associated, not to be used again in the counting sequence (tagging). If an item is not assigned a number name or is assigned more than one number name, the resulting count will be incorrect (illogical). The two processes requiring coordination:
   - The partitioning process (a.k.a., intervaling, pattern recognition) – recognizing [dis]similarity, and the memory/record/awareness its presence. In other words, moving from the to-be-counted category to the counted category.
   - The tagging process (a.k.a., naming, label, neologizing) – identifying, selecting, and assigning a tag, name, or otherwise, label/category. Naming the separation of [dis]similarity. There are at least three tags for every [dis]similar pattern:
   - A tag representing the category of pattern presenting itself as a sequence/count of intervals and patterns. This is generally notated/expressed as a word (number, e.g., two, three) or letter.
   - A tag representing the sequence of the pattern. This is generally notated/expressed as a numeral.
   - A tag representing the interval of the pattern. This is generally notated/expressed as a numeral.
2. **The stable-order principle** – The counted tags must be arranged in a stable (i.e., repeated) order. To be able to count also means knowing that the list of words used must be in a repeatable order. This principle calls for the use of a stable list that is at least as long as the number of items to be counted; if you only know the number names up to 'six', then you obviously are not able to count seven items. For example, someone who counts 1, 2, 3 for one particular collection of three objects, and 2, 1, 3 for a different collection, cannot be said to have an understanding of the stable-order principle – although that person would appear to have an understanding of the one-one principle. However, a person who repeatedly counts a three-item collection as 2, 1, 3 does appear to have grasped the stable-order principle – although, in this case, has not yet learned the conventional sequence of number names.
   - From this principle comes the **radix/base** of all number systems.
3. **The Cardinal principle** - On condition that the one-one and stable-order principles have been followed, the number name allocated to the final object in a collection represents the number of items in that collection, its 'value'. The last number-word of an array of counted items has a special meaning: it represents the set as a whole [value] and the numerosity of this set of items. Note that the cardinal principle pre-assumes the one-one and stable-order principles [are encoded]. The final

number name is different from the earlier ones in that it not only 'names' the final object, signaling the end of the count, but also tells you how many objects have been counted: it indicates what is called, the numerosity of the collection. If someone recounts a collection ({1,2,3,..}) when asked, how many objects there are, then they have not yet grasped this principle.

These three principles are considered by Gelman and Gallistel to be the 'how-to-count' principles as they specify the way in which the counting operation must be executed (i.e., proceed). The remaining two are 'what-to-count' principles, as they define what can actually be counted.

4. **The abstraction principle** – The real-ization of what is counted. The logical mapping of relationships in consciousness through a nominal "scale" (name only) measure stored in memory.

When the how-to-count principles are combined with the abstraction principle, there is an order of magnitude rise in functional expressibility (i.e., enhanced creativity).

5.7 Numeral

A numeral is any symbol or word for a number. A numeral is a symbol or name that stands for a number. The number is an idea (signified), the numeral is how we write it (signifier). That which is being counted, measured, or otherwise given a numerical value is the sign, or otherwise, sensation.

5.8 Numbers and numerals

A number is a specific value (quantity or [ac]count). Other words for the term 'number' include, but may not be limited to:

1. Quantity – how much?
2. Value – what size?
3. Count – what placement?

There are two types of counts:

1. The points that determine the boundaries
2. The spans between the boundaries.

A 'number' is an abstract [theoretical] concept, and a numeral is the way that people denote that concept. A number is an abstraction represented by a symbol called a numeral. A numeral is a symbol or name that stands for a number. What that symbol looks like is technically irrelevant. However, if a meaningful relationship were present between the symbol and the meaning (number) it conveyed, then that would be optimal for efficient processing. Note that frequently the words "number" and "numeral" are used synonymously/interchangeably, although technically, there is a difference.

Numbers express meaning, and numerals are the symbols (signifiers) used to communicate the meaning. Therein, numbering is the assigning of meaning in the form of a number to something. As an ad-jective, 'numerical' means expressed in numbers, or relating to numbers.

A numeral contains one or more written symbols, but a number can be expressed in a range of ways. A number is a concept whose types can be expressed in [at least] word form (i.e., linguistically; e.g., one, two, three) and symbol form (i.e., mathematically; e.g., 1, 2, 3). The different ways of visually representing numbers are referred to as numerical notational forms.

> **NOTE:** *Something for which mathematical logic has no application is not a number. For example, a so-called "telephone number" is not a number. The symbols/digits in a telephone "number" cannot be added together to get another number, or any relevant mathematical pattern. If mathematics cannot be applied, then symbolized identifier (e.g., "telephone number") is not a number. A telephone "number" is a sequence of digits assigned to a communicating user. It is essentially, the name or address of the user. It is a string of decimal digits (i.e., sequence of digits) that uniquely indicates a network termination point (or user), and is required in the routing of network traffic; it is an identifier assigned to a user. A "telephone number" is a data structure, it is not a single value. Similarly, an IP address is not a number as such, but a string representation of X number of bits/bytes. The IPv4 (protocol) address is made up of 32 bits (4 bytes, 4 octets). The IPv6 (protocol) address is 128 bits (16 bytes, 16 octects). A number is not the same as a location identifier, there is a difference of type.*

5.9 Digits in number system

Digits are the finite number of symbols used in a number system. A digit is a single symbol used to make numerals.

- 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9 are the ten digits used in everyday numerals.
- Example: The numeral 153 is made up of 3 digits ("1", "5" and "3").

A digit in base nn is one of the nn symbols used to represent the whole numbers zero through n−1n−1.

For the decimal (base 10) number/numeral system, the digits are: 0,1,2,3,4,5,6,7,8,9 .

Thus, in base ten the digits 0,1,2,3,4,5,6,7,8,9 are used to represent the numbers zero, one, two, three, four, five, six, seven, eight, nine. In base two, binary, the digits are usually 0,10,1 and are used to represent those same ten numbers as 0,1,10,11,100,101,110,111,1000,1001 0,1,10,11,100,101,110,111,1000,1001. In base sixteen, sexadecimal ("hexadecimal"), another six digits are required to represent the numbers ten through fifteen. The digits normally chosen are A,B,C,D,E,FA,B,C,D,E,F.

Digits are used to compose numbers. A digit in mathematics is like a letter in linguistics. A number can be represented by one digit (0-9 are one-digit numbers) or more than one digit (10 and greater are two or more digit numbers). A digit (letter) is one of the individual symbols used in writing a number (word), and the term "numeral" could refer to either one of the symbols or the set of symbols used to represent a number (or concept).

For clarification,

1. Digits make up numerals, and
2. numerals stand for an idea of a number.

A number can be a numeral only, or any combination of numerals and signs. A number can be written with one or more words, letters, numerical digits, or symbols. A number is a concept that has various representations/notations/expression:

- Twenty one or two ten one
- XLII
- 42
- -42
- 1010102
- 2A16
- 7x6

Something made up of digits is not necessarily a number. A number has a numeral value, while digit is just a representation.

5.10 Unnecessary complexity

More noise is generated by the different "number formats" that have been developed over the course of history. A small sample of the different forms include: 2 + ½ = 2½ = 5 / 2 = 25 / 10 = 2.5 = 2 + 2-1. The deci-mals are a form or result of division.

5.10.1 Numbers as Objective and Subjective

Numbers are an abstraction (conceptual) and do not exist in the real (physical) world. However, just because something conceptualized does not exist in the real physical world, does not mean that it is subjective (i.e., disconnected from the physical world). The number "4", as in, 4 of something (e.g., 4 coconuts), is not something in the real physical world, but it also is not subjective (i.e., disconnected from the physical world). There is an experiential relationship between the numerical signifier "4", the [as]signed numerical meaning "four", and the physical sensation [of 4/four] of something [which can be characterized as unique or different than other things]. The number "4" does not exist, but that does not mean any numerical signifier [of meaning] can be assigned to [the experience and conscious count/awareness of] 4 coconuts. In other words, if 4 coconuts are present in front of someone sensing the real [physical] world, then it cannot be logically said that 3 or 5, or any number/count other than 4 (four), are present.

The scientific method also does not exist in the real world, but that does not mean that the scientific method is subjective. The idea that if something does not exist in the real [physical] world, that it is then subjective[ly disconnected from that world], is not valid. Some conceptual abstractions objectively express an existent relationship in the real physical world, and other conceptual abstracts do not express any relationship to, or in, the real [physical] world.

When there is no experiential reference point for a conceptual abstraction, that conception is commonly said to be "subjective" – related to a separate[d] subject, which is dis-connected from the other (or, all other) subjects in a unified object[ive world].

5.10.2 The real number line: visual positioning

The number line is a series of dots (Read: points, degrees, or packets of information) representing iteration, along a single dimension. Between each dotted sub-division there is an interval. The interval has a beginning, commonly known as it's "position". Therein, the interval's value is the duration of its position. Each interval has a position [of beginning]. That value has a numerical sequence [value] and a unit [value].

> **NOTE:** *Counting necessitates units and sequence.*

A real number line (or simply, number line) allows for the visual display of real numbers by associating them with unique points (positions) on a line. The real number associated with a point is called a coordinate.

Numbers can be conceptualized to exist along a one-dimensional continuum known as a number line (where 0, positive, and negative numbers, fractions, and [ir] rational numbers are all possibly present).

"Imaginary" numbers are not just left or right on this 1st dimensional number line; they exist in a whole different dimension. Algebraically, this new dimension has the expression, the square root of negative 1, √-1. The result of combining a number along the 1st dimensional number line and a number along the 2nd dimensional number line is a functional two-dimensional form, a "complex" number.

The conceptualization of 'number' includes this extra dimension, which has a referential association with the 1st dimensional number line.

Using the word "imaginary" as a label for this category of number conceptualization is a horrible decision. This extra dimension allows for the full visualization of the functional expression: f(x) = x2 + 1. The function crosses the x axis in a two-dimensional graph in this other dimension. This is an extra dimension that the conceptualization of numbers posses. It has been misnamed. The name suggests these numbers are not as "real" as counting numbers, which is not accurate. It has been suggested that these numbers should instead be given the name, lateral. From here on, lateral means imaginary.

The poor selection of a name for the categorically named positioning operation is the most significant reason why people don't understand that a negative times negative is positive, or a pure positive "imaginary" times pure positive "imaginary" is negative real number.

Carl Friedrich Gauss, who gave the first clear exposition of complex numbers and due to his contributions to the theory of electromagnetism, the international unit of magnetic induction is called by his name, the gauss, wrote,

"If we call +1, -1, and √-1 had been called direct, inverse and lateral units, instead of positive, negative, and imaginary (or impossible) units, such an obscurity would have been out of the question." (English translation from German)

Gauss suggested that the concept presently/previously known as "negative" (-) should be renamed as 'inverse' (-). Logically, inverse times inverse is direct. Or, positive is forward, and negative is backward. If the operation is inverse, and then inverse again, of the result is the original direction, like backward and then backward is forward.

The new names [for the positioning operation] are as follows:

- +1 = positive one OR direct one [unit]
- -1 = negative one OR inverse one [unit]
- √-1 = imaginary one or lateral one [unit]

Gauss's "imaginary" number name is lateral number (side number). When direct lateral times direct lateral, which is √-1 * √-1 = -1, then -1 * √-1 is inverse lateral, - √-1, then inverse lateral * direct lateral is direct, -√-1 * √-1 = +1. The square root of a negative number can be solved for with the square root of -1 times the negative number, logically creating the expression of lateral movement (i.e., a movement different to the original axis).

There are two ways to perceive/categorize an angle:

1. Degrees, the swivel an observer went through to follow an object. Degrees are the observers viewpoint.
2. Radians, the distance the object moved on its path. Radians are the mover's viewpoint.
3. 'Radians' are used in operational physics formulas, including but not limited to sine, cosine, etc., wherein 'radians' refers to the distance the object moved.

5.11 Mathematical functions

> **INSIGHT:** *The names applied to concepts and objects determine how easily concepts and objects may be understood, integrated, and visualized. The wrong names are likely to confuse understanding and limit thinking.*

Mathematical operations are also called, functions, expressions, number sets, and operations. Physical operations are also called, formulas, functions, and operations. The mathematical operation is a notation for mapping patterns of [information] movement.

5.11.1 Integral function

Integrals are usually described as the inverse of differentiation, finding the area under the curve, and so on. Integrals allow for the 'multiplication' of changing numbers. Take for example, "3 x 4 = 12"; what if one quantity is changing? It is not possible to multiply changing numbers, so the next operation is to integrate (integration), a more complex multiplication operation.

With regular multiplication, it can be assumed that the value of one unit holds for the entire whole. Integration (piece-by-piece) is required when there is change/time. Time/change becomes a series of instants, each with its own value. Add up the instances (i.e., distance moved) on an instant-by-instant basis.

Multiplication is the combining of quantities into a new result. Changing quantities can also be combined into a new result. The multiplication operation allows for the changing of quantities into a new result, but not directly for changing quantities. Here, integration is the idea of combining quantities into a new result. It is possible to integrate ("multiply") length and width to get area. Area is just one way to visualize multiplication. Yet, it is possible to integrate speed and time to get distance, or length. When integration is required, but the integration operation of multiplication is not possible, then a different integration operation is required.

In mathematics, an integral assigns numbers to functions in a way that can describe displacement, area, volume, and other concepts that arise by combining infinitesimal data. Integration is one of the two main operations of calculus, with its inverse, differentiation, being the other. Given a function f of a real variable x and an interval [a, b] of the real line, the definite integral, $F(x) = \int f(x)\,dx$, is defined informally as the signed area of the region in the xy-plane that is bounded by the graph of f, the x-axis and the vertical lines x = a and x = b. The area above the x-axis adds to the total and that below the x-axis subtracts from the total. The operation of integration is the reverse of differentiation.

In calculus, an integral is written in the following notation and expressing the following relationships. This function[al operation] (a.k.a., integral notation) describes a piece-by-piece multiplication of two elements:

Integral notation (in calculus):

- distance = \int speed(t) \ Δt
  - distance = \int speed(t) \ dt
- distance = ∫ speed(t) \ Δt
  - distance = ∫speed(t) \ dt

The relationships:

- The integral sign (s-shaped curve) means we're multiplying things piece-by-piece and adding them together.
- dt represents the particular "piece" of time we're considering. This is called "delta t" (Δ), and is not "d times t".
- t represents the position of dt
- speed(t) represents the value being multiplied by

Clarifications:

- Writing the operation as "∫ speed(t)", with an implicit dt makes it easy to forget that this is a piece-by-piece multiplication of two elements. Like any function[al operation], when it is described, it should always be described fully.
- Presently, the section concerning speed is written as speed(t) • dt, instead of speed(t_dt) • dt. The latter makes it clear we are calculating for "t" at a particular change "dt", and not some global change "t".
- The way the letters are used is confusing. "dt" looks like "d times t" in contrast with every equation you've seen previously.

NOTE: *Integrals multiply changing quantities.*

5.11.2 Quadratic function

A quadratic function is one of the form f(x) = ax2 + bx + c, where a, b, and c are numbers with a not equal to zero. The graph of a quadratic function is a curve called a parabola. Parabolas may open upward or downward and vary in "width" or "steepness", but they all have the same basic "U" shape.

A quadratic function is graphically represented by a parabola with vertex located at the origin, below the x-axis, or above the x-axis. Therefore, a quadratic function may have one, two, or zero roots. Here, 'roots' are also called x-intercepts or zeros.

5.12 Integration and the fundamental operations of the concept, 'number'

The concept, 'number' has the following properties, which represent the fundamental operation[ing] of numbers. A number can have [at least] the following operations repeated on it:

- Counting (sequence[d/ing])
  - E.g., 0,1,2,3,4,5,6,...
- Scaling (order of magnitude)
  - E.g., 1,10,100,1000; 10,20,30,40
- Flipping (inverse)
  - E.g., 1,-1;2,-2;3,-3;4,-4
- Rotating (angle,radian)
  - E.g., 3i;9i;3+4i

The logical conception of a number allows for the technical creation of higher functioning, which requires increasingly "complex" calculation operations.

In part, the purpose of 'number' is to express [the logic of] integration. Multiplication is the understood beginning of [numerical] integration. A number can be broken into units (whole and partial). Then, each unit (piece) can be multiplied (duplicated) by a sequence of iterations, and the results can be added. This is the beginning of [mathematical] logical integration.

The application of integration as "multiplication" results in a unique operation at each consecutive level of conceptualization of the concept, 'number':

1. **The addition operation**: Integers can be added together (integrated) into another number, a piece of data, a result [of calculation].
2. Multiplication operation at integer level of conceptualization: Integers can be repeatedly added together (repeated integration). With integers, multiplication is repeated addition.
3. Multiplication operation at negative number level of conceptualization: with negative numbers, multiplication is flipping.
4. Multiplication operation at "real" number level of conceptualization: with real numbers, multiplication is scaling.
5. Multiplication operation at complex number level of conceptualization: with complex numbers, multiplication is rotating and scaling.

NOTE: *Area is a visual representation (i.e., visualization) of multiplication.*

The conceptual evolution of the addition operation viewed from the perspective of sub-division (or pattern recognition):

1. Addition
   - An additional sub-division/pattern in the sequence
2. Ordered addition
   - Groups of additional sub-divisions/patterns in the sequence.
3. Inverse addition
   - Reverse sub-division/inverse of pattern in the sequence.
4. Repeated addition

- Multiplication – Effect varies by the level of conception of 'number'.

5.12.1 Division & multiplication

TERMINOLOGY: *Procept – the combination of process and concept.*

Here, one is the variable, a unit [of something].

1. **Sequencing** - Creating another one (sensing another pattern, an equal interval) – the sensation or creation of iteration.
2. **Synthesis** - Combining more than one into one (e.g., 1+1) – integrating two of the same categorical values into one [whole] value of the same category, addition. A "sum" is produced. The operation: addition of two similar data points (values). The result: the sum "product" of those values.
    - Integration -- Repeatedly combining more than one into one (1+1+1) has its own notation, 1x3, multiplication. Multiplication is the repeated integration of a sequence value and a [whole] category value, which produces a new information "product". This number has two data points, sequence [value] and category [value]. Multiplication is repeated addition.
    - The operation:
    - The result:
3. **Analysis** - Separating one into more than one (e.g., 1/2) - subdividing one interval [value] into more than one interval [value]. This is divisioning (as "subtraction").
    - Repeatedly subdividing one or more, into one or more is called "division". Division produces a new information product known as a 'quotient'. Division is repeated subtraction [to form a new interval].

Analysis could be viewed as measurement, and synthesis as the integration of a measurement toward greater understanding and more refined models.

Relationships between division and multiplication operations include:

1. Division is the opposite of multiplication.
2. Division as a process is a multidimensional notion.
3. Division expresses the concept that from a whole [number] there is an equal divisioning [number]. Division is the separation of a number into equal parts. Division may be viewed as a form of repeated subtraction from a whole.
4. Division is the operation of repeated sub-divisioning. The result of the division operation is new data about a pre-existing number.
5. To divide is to separate one into more than one, equally or unequally. Division is the repetition of separating one into more than one.
6. Multiplication is a form of repeated addition.
7. Multiplication is repeated addition. The result of the multiplication operation is.
8. The multiplication [function as an] operation asks, How many in all; how many [units] in whole [unit]; how many all together?
9. Multiplication is: factor (group sequence) x factor (group value) = product
10. Addition is the combining of two or more. To multiply is to take one out and duplicate it one, and then, combine the sequence and interval values into a "product".

The repeated summation of one "factor" and another "factor". The word factor could be replaced by any word meaning

5.12.1.1 The division model

The division [function as an] operation asks, How many each; how many groups; what is each share/partition?

There are two models for division:

1. Partition division (also known as partitive, sharing and grouping division) is a way of understanding division in which you divide an amount into a given number of groups. If you are thinking about division this way, then 12 ÷ 3 means 12 things divided evenly among 3 groups, and we wish to know how many is in 1 (each) group.
2. Measurement division (also called repeated subtraction division), is a way of understanding division in which you divide an amount into groups of a given size. If you are thinking about division this way, then 12 ÷ 3 means 12 things divided evenly into groups of 3, and we wish to know how many groups we can make.

5.12.1.2 Multiplicative model

Multiplication is:

- To multiply one number *n* (a multiplicand) by another *m* (a multiplier) means to repeat a multiplicand *n* as an addend *m* times. The result of multiplying is called a product.

The multiplicative model is:

- Product (dividend) is factor (divisor) • factor (quotient)
- Dividend – the number being divided
- Divisor or factor – the number that will divide the dividend exactly.

- When a multiplication fact is known, then a division fact must also be known.
- Divisible – can be divided without a remainder.
- Quotient – the result of division

CLARIFICATION: *To duplicate is to make an "exact" (or as close to) copy; a second copy of the pattern. One might offer the idea that instead of subtract, 'sublicate' (reverse of duplicate) means "to take away".*

There are multiplication and division tables. There is one model for multiplication:

1. Repeated addition is the model for multiplication.
2. Multiplier – the number of sets (of a patter).
3. Multiplicand – the value/amount in each set.
4. Product/result – multiplier multiplied by the multiplicand (i.e., multiplier x multiplicand = product). Here, order in the operation is irrelevant.
5. For example, 4x6 = 6+6+6+6 = 24

5.13 Number system and numeral notation

A numeral system or number system is a mathematical system [for the representation] of numbers. A numeral system (or system of numeration) is a writing system for expressing numbers; that is, a mathematical notation for representing numbers of a given set, using digits or other symbols in a consistent manner. It is the applied logic [of a number system] that allows the symbols "11" to be interpreted as the binary symbol for three, the decimal symbol for eleven, or a symbol for other numbers in different bases.

The terms 'number system' and 'numeral system' are often used synonymously. It may be clearer to present a slight differentiation between the two terms, where the term 'number system' represents the internal logic of the system, and the term 'numeral notation' represents the written expression of the system.

NOTE: *The allowance for reusing numerals (symbols) simplifies arithmetic.*

5.13.1 Radix/base

Every number system has a specific [finite] number of unique digits known as its base/radix (i.e., the number of unique digits in the system).

In mathematical numeral systems, the radix or base is the number of unique digits, including zero. Etymologically, 'radix' is a Latin word for "root". Root can be considered a synonym for base in the arithmetical sense. For example, for the deci-mal system (the most common system in use today, coming from the ten fingers/digits of humans) the radix (base) is ten, because it uses the ten digits from 0 through 9. If numerical representations greater than 9 are required, then a new position is required (10,11,12,13,14,...,99,100,...,999,1000...).

DEFINITION: *The word "base" in mathematics is used to refer to a particular mathematical object that is used as a building block.*

5.13.2 Positionality

Number/numeral systems may be categorized by either positional or non-positional notational encoding logic. There are two primary ways in which numbers can be represented/encoded. They can either be encoded positionally, or encoded without position having meaning (i.e., non-positionally). Positional [numeral system] notation is distinguished from non-positional notation by its use of the same symbol for different "orders of" magnitude (different meanings). For example, the "ones place" (1), "tens place" (10), "hundreds place" (100) – the 1 repeats three time, and means something different each time (one, ten, then hundred). In a positional numeral system, the position of any given digit in a number has [logarithmic] mathematical significance.

5.13.3 Non-positional

Non-positional number/numeral system (non-positional notation) – Characters/digits are position invariant, meaning each character represents the same value regardless of its position. In Roman numerals, for example, the symbol V always means "five", whether it occurs last in a numeral string (e.g., XXV), next to last (XXVI), third from last (XXVII) or fourth from last (XXVIII). In the Roman numeric system, each numeral has a fixed value, rather than representing multiples of the base number (e.g., 10, 100, and so on), according to position. Hence, there is no need for "place keeping" zeros. Notice that in the Roman system, position still has relevancy (e.g., I before V or X indicates one less), but that relevancy has no fixed zero relationship, no positional relevance [relative to the number as a whole].
The unary (base-1, tally marks) numeral system, frequently used for counting, is non-position: /=1; //=2; ///=3; ////=4; /////=5; //////=6; etc.

NOTE: *Arithmetic operations are possible, but more difficult.*

Non-positional number systems have a base number of repeating digits, which may be 1 or more. However, they are not (generally) categorically named after their base, they are given cultural names.

Note: It could be said that there are cultural number/numeral systems. Cultural numeral systems involve unique character/symbol visualizations, and they include but are not limited to: Babylonian, Egyptian, Vedic, Greek, Roman, Chinese, Arabic, Hebrew, Indian, etc. In this category, the numerals and their rules are viewed as having arrived due to unique cultural values and symbols. Simply, a cultural numeral system is the name of any given numeral system a specific "culture" uses, and it may be positional or non-positional.

5.13.4 Positional

Positional number/numeral system (positional notation, place-value notation) - Place (position) has value (meaning). Where a digit occurs in a number (as a string of digits) determines its meaning. A positional number system gives different meaning to the same symbol depending on its position. The position dictates rules to manipulate the symbols, not their value (magnitude). In a positional number system, the value of each digit is determined by which place it appears in the full number. A positional (numeral) system is a system for representation of numbers by an ordered set of numerals symbols (called digits) in which the value of a numeral symbol depends on its position. For each position a unique symbol or a limited set of symbols is used.

The base of a positional number system is, how many digits (symbols) there are for each position in a number.

In a positional system, the value of a symbol is given by the order of its position expressed in the bases (or radices) of the system. The total value of the represented number in a positional number is the sum of the values assigned to the symbols of all positions. For each position that the number is in, in that system has a relative symbol or meaning, and in a way relates to the number directly next to it. The total value of a positional number is the total of the resultant values of all positions. For each position that the number is in, in that system has a relative symbol or meaning, and in a way relates to the number directly next to it.

In a positional [notation] numeral system each position is related to the next by a constant multiplier (i.e., base) of that numeral system. The different numeral systems sub-categorized by different bases are given the suffixes -ary, -imal, and -al. Each position represents a different base.

Positional number systems are categorized by their radix/base. Each iteration of the base forms a magnitude of sequentially iterating order (known as "orders of magnitude").

There are generally considered two sets of rules for encoding positional information: [a] "standard" [set of] rules; and [a] "non-standard" [set of] rules:

1. Standard positional numeral systems/notation – whole number orders of magnitude from base/radix 2 onward.
   A. Binary, ternary, quaternary, quinary,...,decimal, sexagesimal,...
   B. The non-standard positional numeral systems are.
2. Bijective numeration
   A. Signed-digit representation
   B. Negative bases
   C. Complex bases
   D. Non-integer bases

In any standard positional numeral system, the number x and its base y are conventionally written as (x) y, although for base ten the subscript is usually assumed and not written, as it is the most common way to express value (by our organism, because of our 10 fingers). For example, (100)10 (in the decimal system) represents the number one hundred, while (100)2 (in the binary system with base 2) represents the number four.

The radix point

With the use of a radix point ("."; e.g., decimal point in base-10), the positional notation can be extended to include fractions and the numeric expansions of numbers into rational and real categories (i.e., into a "real" [one dimensional, root = base/radix] set). Note that the point/dot takes on the name of the numeral system. For example, the point/dot in the:

1. Binary numeral system may be called, a binary point.
2. Quinary system it may be called, a quinary point.
3. Decimal system, it may be called, a decimal point.

5.13.5 Scale

> **NOTE:** *In order to communicate results unambiguously it is necessary for each of us to share the same scale for a quantity and to have access to the standards that define the scale.*

A scale is a totally ordered numerical structure onto which physical quantities are mapped, with the mapping preserving the structure of the original physical quantity. Every scale can be realized using mathematical concepts.

Length and angle scales are realized with mathematical concepts. Hardness, temperature, and other environmental scales are realized using concepts from physics rather than mathematics. Both length and angle scales are linear scales. Length requires a unit length to be defined. Early unit length standards were realized as end bars,in which the fundamental unit was the distance between the two ends of the bar. The meter, the fundamental unit of length, is now defined from concepts in physics, rather than being based on the results of a mathematical survey of a geographical artefact.

The fundamental unit is then divided upon a continuum, into a 'scale'. It may be divided equally, or unequally (and its visual representation as equal or unequal depends upon the particular notation). Mathematically, division is interpolation, which is the computation of new data points from known discrete points. For the length scale, this consists of dividing the unit length, defined by a line scale, into equal lengths, and for angular scales, dividing the circle into equal angles.

Euclid's Elements provides various geometrical constructions to divide lengths into equal parts and constructions for particular angles and angular bisection.

> **NOTE:** *A 'chart' is a two-axis scale (or continuum). A three-dimensional 'Cartesian coordinate system/scale' is a three-axis scale (e.g., x,y,z).*

A scale is a system[s approach] that we use to perceive [existence], by the method of arranging data [of a similar pattern] in [logical] order.

Every scale requires the following elements:

1. Visually expressed with iterative markings representing the division of a pattern.
2. An ordered numerical/linguistic notation.
3. At least one categorical dimension.

The word scale has several applications, all of which relate to the idea that there exists a divisional iteration of a pattern.

> **CLARIFICATION:** *The term 'continuum' has a relationship to the concept of 'scale'. A continuum is a set of iterations on a scale, which have a particular characteristic to different degrees. Any continuous whole comprising of individual units with a logical progression can be considered a continuum. The term is also given to a body that can be continually sub-divided into infinitesimal elements with properties being those of the bulk material. And, a continuum is a region of filled space.*

5.13.5.1 A 'scale' as a tool for understanding a discontinuous category - scale as a discontinuous category (understanding)

There exists a pattern, and the pattern repeats along an ascending or descending scale. Here, a scale is a conceptual-mathematical visualizing tool for divisional categorization. Numbers can be placed/positioned in order (as intervals) along one or more lines to create a visual scale, which represents a discontinuous category of information. The numerical space between divisions may be equal or unequal.

A number/numerical scale is a line on which the marks of separation have been given numerical names or labels. A number scale is constructed by starting with a line, and a line segment of fixed amount (or magnitude), which represents the first level (or order) of magnitude. Levels (or orders) therefrom may exist at equal or unequal separated magnitudes.

A 'scale' is a way of visualizing spatial size/quantity categorizations with division marks indicating (divisioning, sectioning, or proportioning). Here, a scale is a type of discontinuous, ordered rank of categories or sizes. More generally, a 'scale' refers to a differentiated category of size. Here, the term 'scale' implies the discontinuous (divisioning) idea of, orders of magnitude. A scale forms a sequence of ascending or descending units/intervals of equal or unequal proportion.

> **NOTE:** *A scale can be used as an organizing structure for understanding something.*

A scale mathematically and/or visually represents portions of a whole. A proportion is a way of expressing how the size or magnitude of one thing relates to that of another.

The degree of separation between iterations (e.g., adaptive repetition of 0-9 (1st order of magnitude), to 10-99 (2nd order of magnitude), and so on) may be conceived of as a number, named a 'radix' or 'base'. There are ten [conceptual] degrees of separation between each iteration in the deci-mal number/numeral system (0-9). When these iterations are expressly notated along a [number] line, that is called a number scale.

Each iteration [of the pattern] on the scale (1,10,100,...) represents a positionally relevant ["order of"] magnitude based on the base/radix of the applied number/numeral system. The presence of an ordered difference (rank) among a set of something similar is embodied by the word "order", as in, "order of magnitude".

The order of magnitude [of a scale] is dependent on its base (radix), and on whether the scale is linear or non-linear.

A scale may have equal or unequal divisioning. A non-linear scale consists of unequally spaced divisions (sections or proportions). A scale that has equal divisions is called 'linear', and an unequally divided scale is called 'non-linear':

1. **Linear scale (equally divisioned scale, proportional scale)** – where the divisions (marks) are evenly spaced. On a linear scale, a change between two values is perceived on the basis of the difference between the values. For example, a change from 1 to 2 would be perceived as the same amount of increase as from 4 to 5. Visually, each line, grid or marking [visible on the scale] is equal in value or size. The divisions, sections or proportions on a linear scale are directly proportional (i.e., equal). The relationship between the variables is directly proportional. Thus, a linear scale is sometimes called a, **proportional scale**. A proportional/linear scale always has an "order of magnitude" [off difference between intervals] of one.
   A. As a "gauging/comparing" tool, the linear scale is used to obtain the accurate measurement of: distance, mass, volume, etc.
   B. Machines utilize a linear scale in order to produce precisely desired outputs. Examples of linear scale tools include: ruler, measuring tape, measuring cylinder, graph sheet, etc.
2. **Non-linear scale (unequally divisioned scale, non-proportional scale)** – where the divisions

(marks) are unevenly (or not equally) spaced. The relationship between the variables is not directly proportional. In a non-linear scale, the divisions, sections, or proportions are uneven/unequal. This means that the visible lines, grids, or other divisional markings (which may appear equally spaced in the visualization) are not equal or constant in value or size. The divisions, sections, or proportions are not directly proportional. Thus, a non-linear scale is sometimes called a, **non-proportional scale**. Note that because the divisions (marks) are not evenly/equally spaced, it is more challenging for a human to accurately read the scale.

A. The **logarithmic scale (order of magnitude scale)** is a well-known type of non-linear scale. Visually, each mark on the log scale is the previous mark multiplied by a value. In a logarithmic scale, values are proportional to the logarithms of the scale numbers. On a logarithmic scale, a change between two values is perceived on the basis of the ratio of the two values. That is, a change from 1 to 2 (ratio of 1:2) would be perceived as the same amount of increase as a change from 4 to 8 (also a ratio of 1:2). A logarithmic scale implies and is based on "orders of" magnitude, rather than individual incrementation, as in a linear scale. Each mark on the logarithmic scale is calculated to be the previous mark multiplied by a value set for the log. A logarithmic scale is marked off in orders of magnitude, that is, each mark on the scale as you move left to right is larger by a multiple of the scales set value, than the one preceding it. If the scales value is 10, then one mark to the right is 10 times larger and one mark to the left is 10 times smaller. On a linear scale the distance from 1.00 to 10.0 is ten times longer than the distance from 0.1 to 1.0. On the logarithmic scale these two distances are equal. Take note that logarithmic does not always mean base 10.

B. The decade log scale is one of the most well-known log scales. One decade is a factor of 10 difference between two numbers (an order of magnitude difference) measured on a logarithmic scale. A decade is a set of ten, or an interval of ten. In the decade log scale, there is a base 10 interval between increasingly higher order whole-number exponentials. For example, 100(1), 101(10), 102(100), 103(1000, 1k), 104(10000, 10k), 105(100000, 100k).

C. One of the most well-known ratios is the "golden" ratio (phi). In geometry, a golden spiral is a logarithmic spiral whose growth factor is φ, the golden ratio. The Golden ratio is a special number found by dividing a line into two parts so that the longer part divided by the smaller part is also equal to the whole length divided by the longer part. It is often symbolized using phi, after the 21st letter of the Greek alphabet. The term 'Phi' was given to the ratio number "in honor of Phidias, the lead sculptor of the Parthenon in Greece". In an equation form, it looks like this: a/b = (a+b)/a = 1.6180339887498948420 ...

> **CLARIFICATION:** *In a non-linear system a change in the output is not proportional to a change in the input. In a linear system, a change in the output is proportional to a change in the input.*

5.13.5.2 A 'scale' as a tool for comparing (gauging)

A scale is a tool used to compare new information to a pre-existing [dimensional] iterative pattern of information, producing new data. A "scale" often signifies a receptor or method that can reliably map a number to a given phenomenon. In this usage, the term may also be part of the proper name of the method signified. This usage implies a comparison of entities, but not the discontinuous idea of "orders of magnitude".

As conceptual instruments, for example: the Mohs scale (hardness), Scoville scale (heat of capsaicin), Kelvin scale (temperature), ph scale, Borg scale (physical exertion), Richter scale (earthquake), and stellar magnitude.

As a physical tool (instrument), confusingly, a scale is also the name of a measuring instrument for weight, as well as another name for a ruler (an instrument for measuring length). Common [physical] length scale measurement gauges include: the Vernier scale, linear scale, engineer's scale, architects scale, scale of duration, scale of calibrated dial.

> **CLARIFICATION:** *Measurement display "gauges" are instruments most often used in situations where the thing being measured changes regularly in time, such as in the measuring of volume of something we being used (e.g., fuel gauge).*

5.13.5.3 Scale in mathematics (fractioning)

"Scale" is a common term in mathematics, usually signifying a proportion. This usage implies a comparison of entities, but not the discontinuous idea of "orders of magnitude".

5.13.5.4 Scale as continuous representation system (modeling)

"Scale" can refer to a continuous representation system (model) that signifies an inter-related set of phenomena.

This usage implies a comparison of entities, but not the discontinuous idea of "orders of magnitude". One thing represents another thing, but at a different size. The 'scale ratio' of a model represents the proportional ratio of a linear dimension of the model to the same feature of the original.

For example, a smaller 3-dimensional "scale" model of a building, or the scale[d down] drawings of the elevations or plans of a building.

The scale can be expressed in four ways: in words (a lexical scale), as a ratio, as a fraction and as a graphical (bar) scale.

- Lexically - One centimeter to one meter
- Ratio – 1:100
- Fraction – 1/100

To scale something is to produce a smaller or larger representation of something. The scaled [down] version of the thing is somewhat confusingly called a 'model' [representation]. A "scale model" is a [physical] model, a representation or copy of an object that is larger or smaller than the actual size of the object, which seeks to maintain the relative proportions (the scale factor) of the physical size of the original object. Very often the scale model is smaller than the original and used as a guide to making the object in full size.

Other examples of a scaled representation system include: the scale on a topological map, a musical scale, and a gauge of measurement.

An object (or representation) can be scaled proportionally and non-proportionally.

1. If an object is being scaled, and its representation maintains proportions after scaling, then it is a proportional scale (i.e., the scaling process/operation used a fixed ratio).
2. If an object is being scaled, and its representation does not maintain proportions, then it is a non-proportional scale (e.g., it may have been scaled along the x-axis, and not equally along the y-axis). This type of scaling is sometimes known as sub-dimensional scaling (i.e., scaling of the unique dimensions of some thing).

5.13.5.5 Scale as providing numerical measurement

Scaling is a term used to describe the way that an operational definition can be conceptualized to provide numerical measurement. Usually the term is applied only to ordinal or interval level measures, as nominal scaling is really just a matter of classification within a set of categories, as we saw above. There are a vast number of different scaling techniques and procedures.

1. **Counting frequencies** - the simplest scaling involves natural measures like the counting of instances of occurrence of events. Such occurrence is absolute in nature and can be measured in terms of its "frequency". Scales reflecting measures of frequency are at the ratio level of measurement.
2. **Measuring magnitude** - of which the Likert scale is a typical example. In this measurement procedure, verbal "anchors", which define the extremes of the dimension being measured, are provided to allow a range of responses to some specific question. It is a mistake to assume that the measurement obtained from magnitude scales such as the ones above is at the interval or ratio level because we have no way of determining that the distances between adjacent scale points are really equal.

5.13.6 Order of magnitude

The term/phrase 'order of magnitude' is used to mean more than one thing. The term 'order of magnitude' has two meanings. In its first meaning, it refers to a type of scale. In its second meaning, it refers to a degree (or iterative mark) in a scale of the 'order of magnitude' class. Here, orders of magnitude also known as degrees of separation, and an 'order of magnitude' is one degree of separation, one interval (one sequence).

5.13.6.1 In concern to the term as a class of scale

An order of magnitude is the class of scale (or magnitude) of any amount, where each class contains values of a fixed ratio to the class preceding it. In other words, an order of magnitude is a scale of repeating numerals with a fixed multiple factor (ratio). Here, the term 'ratio' is the relative magnitudes of two quantities (usually expressed as a quotient (the result of division).

In a scale of the 'order of magnitude' type, there is an exponential change of plus-or-minus 1 in the value of a quantity or unit along the continuum.

Any whole number can be an order of magnitude, because any whole number can be radix/base.

In a linear scale, the fixed ratio is one. In a non-linear scale the fixed ratio is not one. An order of magnitude is a number assigned to the ratio of quantities. If the ratio of quantities at each interval is one, then a linear scale exists. If it is not one, then a non-linear scale exists.

Mathematically, the logarithmic scale is used to calculate orders of magnitude.

If the amount being scaled is 10, and the scale is the base 10 exponent being applied to this amount, then to be an order of magnitude greater is to be 10x (times) as large. Such differences in order of magnitude can be measured on the logarithmic scale in "decades" (i.e. factors of ten). Therein, if there are two quantities are of the same order of magnitude, and if one is less than 10 times as large as the other, then the number of magnitudes that the quantities differ is specified to within a power of 10.

The 'order of magnitude' of a scale is the constant factor (ratio) used in division or multiplication to increment a value on the scale.

1. "One order of magnitude more than a given value" - means the multiplication of a given value by the factor (a.k.a., power) of the scale. 100 is an order of magnitude larger than 10. (102 > 101)
2. "One order of magnitude less than a given value" - means the division of a given value by the factor (a.k.a., power) of the scale.

When the ratio/factor is 10, then:

1. One order of magnitude more than 1, is 10 (101); and, one order of magnitude less than 1, is 0.1.
2. Six orders of magnitude more than 1 is 1,000,000 (a million or 106).
3. A value growing by four orders of magnitude implies it has grown by a factor of 10,000 or 104.
4. The order of magnitude of a final number is the number of powers of 10 contained in the number. The number of powers of 10 contained in 10000 is 104.

5.13.6.2 In concern to estimation

Order of magnitude means a number's nearest power [of some base]. If the magnitude of order is 10, then this means a number's nearest power of ten.

5.13.6.3 In concern to the term as a degree in an 'order of magnitude' scale

An order of magnitude is a degree, or a degree change, in a continuum of size or quantity (of measurement). Here, the term 'magnitude' is the property of relative size or extent (whether large or small).

- "Its length was on the order of a meter".
- "The explosion is of a low order of magnitude."

5.13.7 Exponents

When a number is multiplied by itself (e.g., 2x2, 4x4, 10x10), the process is called squaring. When a number is multiplied by itself three times (e.g., 2x2x2, 4x4x4, 10x10x10), the process is called cubing. A number multiplied by itself four times has no unique name/label, and is, and thereafter, "raising it to the fourth (fifth, sixth, ...) power". Squaring is raising to the second power, and cubing is raising to the third power.

The power to which a number is raised is the exponent of that number:

- $base^{exponent}$
- $base^{power}$

There are two commonly accepted notations for the mathematical operation of "raising to a power". For example, raising ten to the power of two:

1. 10^2
2. 10^2

A number can be raised to any power, including decimals. The logarithm of a number is the power that some base number must be raised to get that number.

Logarithms compress scales. A linear scale is like a ruler on which each step on the scale adds a unit: to get two meters, one meter is added to one meter; to get three meters, one meter is added onto another, to which another is added. Conversely, on a logarithmic scale, each step on the scale is a multiple of the preceding step.

For example,

- Log10 1 = 0
- Log10 10 = 1
- Log10 100 = 2
- Log10 1000 = 3
- ...

If the logarithms of two numbers are added together, the result is the logarithm of the product (not the sum) of the two numbers. This reflects the fact that steps on a logarithmic scale are multiples.

When a number is multiplied by itself more than once, it can be expressed (in notation) in terms of an "exponent" - the exponent is a little number to the upper right of the number that says, "this is how many times the number has been multiplied by itself". So,2 x 2 x 2 = 23. The "logarithm" is the reverse of this operation. When we ask, "what is log_2(8)" we are asking, "what is the base 2 logarithm of the number 8", or, "how many times did we multiply 2 (the base) to get the number 8". The answer to this question is the exponent from above.

Logarithms are useful in comparing values that vary over a large range.

In mathematics, the logarithm is the inverse operation to exponentiation. That means the logarithm of a number is the exponent to which another fixed number, the base, must be raised to produce that number. In simple cases the logarithm counts factors in multiplication. For example, the base 10 logarithm of 1000 is 3, as 10 to the power 3 is 1000 (1000 = 10 × 10 × 10 = 103); 10 is used as a factor three times. More generally, exponentiation allows any positive real number to be raised to any real power, always producing a positive result, so the logarithm can be calculated for any two positive real numbers b and x where b is not equal to 1. The logarithm of x to base b, denoted logb(x), is the unique real number y such that by = x. For example, log2 64 = 6, as 64 = 26.

- 26 = 64
- 2 is the base
- 6 is the exponent
- 64 is the result of the operation
- Log2 64 = 6 or Logbase x = exponent

On a logarithmic scale, each delineation/division ("tick mark") on the scale is the previous tick mark multiplied by some number (or value). A logarithmic scale is a nonlinear scale used when there is a large range of quantities. It is based on orders of magnitude, rather than a standard linear scale with equal divisions.

A **physical logarithmic scale** is a scale [of measurement], a tool for comparison, that uses the logarithm of a physical quantity instead of the quantity itself.

Exponential logarithmic notation:

Table 40. Measurement > Numbers*: Table showing base 10 counting in exponential and logarithmic form.*

Exponential form	Logarithmic form
10^3=1000	Log10(1000)=Log1000=3
10 is base	

5.13.8 Common positional number/numeral systems

The most common include:

1. **Unary (non-positional)** – every natural number is represented by a corresponding number of symbols. If the symbol "/" is chosen, for example, the number seven would be represented by /////// (seven of the symbol "/"). Any number (i.e., any value) can be represented by combining these digits. The unary [numeral] system can be modified by introducing different symbols for certain values.
2. **Binary (positional)** – two digits (or numerals), 0 or 1. Any number (i.e., any value) can be represented by combining these two digits. This is a base 2 (binary numeral) system. Hence, there are two values. The binary numeral system can be physically implemented with a two-state device.
    - Positional systems obtained by grouping binary digits by three (octal numeral system) or four (hexadecimal numeral system) are commonly used.
3. **Decimal (positional) Arithmetic [numeral system]** – Decimal representation refers exclusively, in common use, to the written numeral system employing numerals as the digits for a radix 10 ("decimal") positional notation. The ten base digits (or numerals): 0,1,2,3,4,5,6,7,8,9 or (0,...,9). Any number (i.e., any value) can be represented by combining these digits. This numeral system is sometimes confusingly called the "arithmetic numeral system". The value assigned to a digit is applied/processed positionally: one's place (1), ten's place (10), hundred's place (100). The system is composed of ten digits, and hence, the position of a digit is used to signify the power of ten that the digit is to be multiplied with: 304 is equivalent to (=) 3×100 + 0×10 + 4×1; or more precisely 3×102 + 0×101 + 4×100.
4. Phi numeral system (positional) – The phi numeral system is also known as: golden ratio base, golden section base, golden mean base, phi-base, base-ɸ, and phinary. It uses the "golden" ratio (symbolized by the Greek letter ɸ, the irrational number (1 + √5)/2 ≈ 1.61803399 symbolized by the Greek letter φ) as its base.
    A. Additive systems - In additive systems numbers are formed by putting together (in a row) several single characters in order of descending value with each character being repeated as many times as required. In expression, this type of system is known as unary/additive notation. Note that additive systems may have additive and subtractive notation. For example, the best known form of additive notation is the Roman system which was similar to the ancient Greek system using letter symbols for powers of 10 and for the intermediate numbers 5, 50 and 500. The symbols used were I for 1, V for 5, X for 10, L for 50, C for 100, D for 500 and M for 1000. Thus 1969 would be written as MDCCCCLXVIIII. A subtractive notation was also used so that, for example, 4 could be written as IV as well as IIII, and 1949 as MDCCCCXLVIIII.
    B. Multiplicative systems - In multiplicative systems there are two kinds of symbols with the symbols of one kind modifying multiplicatively the values of the second kind of symbols.
    C. Arithmetic table

A numeral is a symbol or group of symbols, or a word in a natural language that represents a number. A numeral system (or system of numeration) is a way to write numbers. Roman numerals and tally marks are examples. "11" usually means eleven, but if the numeral system is binary, then "11" means three. Bases categorize numeral systems?

The **decimal point** is the dot (.) placed after the figure representing units in a decimal fraction. A decimal mark is any symbol used to separate the fractional part of a decimal from the whole part. A decimal number usually means there is a decimal point (.) in the number. The decimal point is exactly to the right of the units position and sets the reference standard for all other positions. The number to the left of the decimal is called the "whole number".

For example,

- 17.591
    - 17 is the whole number.
    - Every movement of a digit further left gets 10 times bigger.
    - Every movement of a digit further right gets 10 times smaller.

- 0.1 = 1/10 = 1 tenth
- 17.591 = 17 + 5/10 + 9/100 + 1/1000

5.13.8.4 Number system bases

This is just a way of writing a value down.

1. Roman numerals: I (1), V (5), x (10), L (50), C (100), D (500), M (1000)
2. Base 10 is a number system that uses 10 digits: 0-9.
3. Base 2:
4. For bases bigger than 10, capital letters are used as symbols. For example, the sexadecimal (a.k.a., hexadecimal) numeral system (base 16) uses the numerical digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F.

The most common bases are binary and hexadecimal (used by computers) and decimal (used by people, because of their ten fingers).

5.13.8.5 Converting [between] bases

In order to convert a decimal number into its representation in a different number base, we have to be able to express the number in terms of powers of the other base. For example, if we wish to convert the decimal number 100 to base 4, we must figure out how to express 100 as the sum of powers of 4.

- 100= (1 * 64) + (2 * 16) + (1 * 4) + (0 * 1)=
- (1 * 4^3) + (2 * 4^2) + (1 * 4^1) + (0 * 4^0)

Then we use the coefficients of the powers of 4 to form the number as represented in base 4:

- 100=1 2 1 0 base 4

5.13.8.6 Numeral system sub-inputs

A numeral system has the following sub-inputs:

1. The symbols: Roman numerals, binary, decimal, fractions, scientific notation, etc.
   - For example, the decimal system has the following symbols: 0,1,2,3,4,5,6,7,8,9, -
2. Rules for combining
   - For example, the decimal system has the following rules:
   - Ordering: -9,-8,-7,-6,-5,-4,-3,-2,-1,0,1,2,3,4,5,6,7,8,9
   - Every symbol to the right represents a one value (or count) increase [in the expressed quantity].
   - Combining: 11,12,13,23,33,45,...
   - When a symbol appears to the right of another symbol it is added.
   - Scaling: 10,20,30,100,200,300,1000, 1100,...
   - Every additional digit represents an increasing (left) or decreasing (right) factor (a.k.a., multiple) of ten.
3. Thus, the creation of a logical numeral system: 0,1,2 ,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,30,40, 50,60,70,80,90,100,200,300,...

Numeral systems may be classified by the symbol (and therein, the number of symbols used):

Table 41. __Measurement > Numbers__*: Table shows base 10 counting symbols in different languages.*

Name	Base number of symbols	Symbols
Greek numerals	10	ō α β γ δ ε ζ η θ ι ...
Western Arabic numerals	10	0 1 2 3 4 5 6 7 8 9
Roman	10	N I II III IV V VI VII VIII IX XL C D M

5.14 Classifying numbers

Numbers can be classified into sets, called 'number systems'. Numbers can be divided into different collections.

5.14.1 Common number types

The concept, 'number', has a three-attribute categorization. There are three common number types (attributes), which have been given the names: nominal, ordinal, and cardinal. In other words, ordinal, cardinal, and nominal numbers are the three [common] attributes of the concept, 'number'.

Cardinal numbers are also known as "counting numbers" and are used to count things. Cardinal numbers are the symbol-unit response to the [numerical] inquiry, "How many?". Ordinal numbers are also known as "position numbers" and are used to place things. Ordinal numbers are the symbol-unit response to the [numerical] inquiry, "Which one?" (or, "What position?").

If there is only the number sequence of set S = {1, 2, 3, ...}. Then, cardinal number of set A" is the number of elements in the set. It is not a specific kind of number (like rational or complex). Similarly, ordinal is not a special number, but merely S applied to ordering. The following question is meaningless, "Is zero an ordinal or cardinal number?" Zero can be the value of the cardinal number of a set. Whether counting starts with 0 is an issue of convenience, though not entirely logical. When you have a list of elements, it is not so practical to start the labeling with 0, since the rank numbers might become adjectives that differ from the proper ranks. However, the tendency would be to associated "level 3" with "the third level", with "third" the adjective of "three". It appears difficult to suppress that tendency. Hence it is better to start lists with label 1.

1. **Ordinal (expresses position)** – order/place (e.g., first, second, third, etc.)

- Ordinal measures refer to the order or the measure, like the order of the cardinality.

2. **Cardinal (expresses iteration)** – pattern/similarity; whole numbers (e.g., 1,2,3,etc.)
    - Cardinal measures refers to the size of something, “How large?”. The interval ‘level/scale of measurement’ is also known as the cardinal level of measure.

Cardinal and ordinal numbers can be used together in the same argument (i.e., in the same sentence).

Nominal numbers are also known as “categorical numbers” and are used to categorize things. Nominal/categorical numbers are numeric codes - numerals used for labelling or identification only.

- **Nominal (express category)** – structure through name (naming) and identity (identifying). Nominal numbers can be single (e.g., 2, 4, 5, 3, 1) or grouped. (e.g., 234, 4432, 53, 3344, 153).

5.15 Number-type scales

There are two types of scales, pure scales and compound scales. A bivariate responses with one response ordinal and the other continuous is an example of compound scales. For pure scales, there are several types:

1. Nominal
2. Ordinal
3. Interval
4. Ratio

5.16 Number

Numbers are strings of digits used to indicate magnitude. One number is the designation of a quantity. For clarification, it is useful to have different words for the two expression of the concept ‘number’. For example, the Dutch language has different words for number (“getal”, as in the list of natural numbers, or the pure decimal system, old-English “tale”) and cardinal number (“aantal”, the number of elements, English “tally”). Historically, the concept ‘number’ was synonymous, but it has since been given a broader meaning (i.e., negative, and complex numbers). Hence, in Dutch, the broader meaning of number is called getal, and aantal refers to the cardinal number or count [of something]. Aantal is an arithmetic value, expressed by a word, symbol, or figure, representing a particular quantity used in counting and making calculations, and for showing order.

In measurement applications, numbers measure the presence of a quantity, known as ‘magnitude’. There are multiple types of numbers, which fall into two principal categories: counting numbers and scalar numbers.

1. Counting numbers (a.k.a., natural numbers, whole numbers, finite cardinal numbers) – count the presence of something. Positive whole numbers, which have no fractional parts. There are no negative counting numbers. Counting numbers stop at zero.
2. Scalar numbers - measure some quantity to any desired degree of accuracy.

Math can be applied to both counting and scalar numbers. For counting, 2 apples + 3 apples = [a count of] 5 apples:

Table 42. __Measurement > Numbers__*: Table shows counting and scalar number ordering.*

Set A is counted using ordered {1,2,3...}	Order in A is not relevant	Order in A is relevant
Counting (process; “order some or all”)	{1,2,3...}	{1st, 2nd, 3rd, ...}
Cardinal (result) (“how many elements are there?”)	{1,2,3,...}	{1,2,3,...}

5.16.1 Numbers as presence and absence

There are *two principle types of numbers* categorized by [the] presence [of existence].

1. The non-zero numbers (presence) - A non-zero number can be used for two purposes: to describe the size of a set, or to describe the position of an element in a sequence. In any number system there is must be more than one symbol used to represent the concept of presence – presence cannot exist without a relationship indicating the presence of two things.
    - For example, the symbols (digits): 1,2,3,4,5, etc.
2. The zero number (absence) - A zero number is used for the absence of a set. In any number system there is only one symbol used to represent the concept of whole absence.
    - For example, the symbol (digit): 0

5.16.2 Numerals used as identifiers and not numbers

The numeral symbols (numerals) are often used as identifiers. Instead of measuring the magnitude of something, or counting some things, these symbols are used to label [objects/events in the real world]. For example, a licensing identification (ID) “number” is not a number; it cannot be used to measure anything, and mathematics does not apply. It is simply a string of symbols/characters that identifies one particular ID from many IDs. Arithmetic cannot be done on the IDs because they are not numbers, they are identifiers (i.e., labels). For efficiency, the selection of identifiers should make logical [conceptual].

Identifying “numbers” (i.e., numeral identifiers) are neither counting nor scalar numbers; instead, they are the symbols used to identify something or act as an identifying label. For example, a phone "number", or id "number" are not scalars.

5.16.2.1 Cardinal numbers (a.k.a., whole numbers, natural numbers, or counting numbers)

In mathematics, cardinal numbers, or cardinals for short, are a generalization of the natural numbers used to measure the cardinality (size or magnitude) of sets. The cardinality of a finite set is a natural number: the number (count) of elements in the set. Cardinal numbers are the natural numbers beginning with 0. The counting numbers are exactly what can be defined formally as the [finite] cardinal numbers.

> **NOTE:** *The transfinite cardinal numbers describe the sizes of infinite sets.*

When we have a set of objects, the cardinality of the set is the number of objects it contains.

The scale of cardinal numbers are (i.e., the cardinal number scale is):

- 0,1,2,3,4,5,6,7,8,9,10,...

Cardinal numbers are integers that can be zero or positive. The usage of a cardinal number assume that the thing(s) being counted are not divisible. There can be 4 of a system, but never 3½ of a system.

Formally, counting numbers are the set of all non-negative integers.

5.17 Number notation

There are various ways that numbers can be written or diagrammed:

1. The number line – a number line is a graphical way to visualize numbers by placing them on a straight line, usually with zero in the middle, positive numbers to the right and negative numbers to the left.
2. Decimal notation (a.k.a., decimal notation) - The most common way to represent real numbers. A string of digits and a decimal point (dot). Digits to the left of the point are increasing powers of ten, those to right are increasing negative powers of ten. For example, 456.65 and -385.109. The numbers on the left side of the dot represent whole numbers, and the numbers on the right side represent decimal values. The point/dot is a decimal signifier – signifying that the numbers coming after it (to the right) are decimal (and not whole) numbers.
   - Note: Different countries officially designate different symbols for the decimal point. In most English-speaking countries, the decimal point is usually denoted by a period/dot to separate the whole number from its fractional parts. However, in continental Europe, the decimal point is usually denoted with the comma. The choice of symbol for the decimal point affects the choice of symbol for the thousands separator, which is largely used in digit grouping.
   - Note: In computing, dot-decimal notation is a string of digits of decimal numbers, each pair separated by a full stop (dot). For example, 192.168.0.1 or 255.255.255.0. The dot in computing is always represented as a dot and never as a comma.
3. Ratios/fractions – a fraction is two quantities written one after the other with a symbol indicating that one is a ratio (or fraction) of the other. For example, 3/4 [of an apple].
4. Normal form (scientific notation) - a number in normal form consists of two parts: a coefficient and an exponent (power of ten). For example, the distance to the sun is 93000000 miles. This can be more conveniently written as 93×106 miles. 93 is the coefficient and 6 is the exponent.

5.17.1 Whole numbers

While the English number words from 11-100 undergo sound changes, the Chinese numbers remain predictable.

5.17.2 Decimal numbers

Decimal numbers are read with the whole number read first, the word dian3 (to denote the decimal point), and then each decimal place read.

> **NOTE:** *Decimal point, decimal, point, and the word "and" can all be used to express the presence of a decimal point.*

5.17.3 Signed numbers

Signed Numbers are the numbers that need an additional explanation to be determined. Usually we have quantities that can be determined with arithmetic numbers. For example: the road length between two cities (for example: 124 miles), the weight of a bottle of milk (for example: 1 Gallon), the land area (for example: 1000 sq. ft.), ... Those are well determined numbers.

Other times we have quantities that need additional explanation, besides the arithmetic numbers, to be determined. For example if we need to drive 50 miles from the point we are, we need to know in what direction; when we talk about time from the present, we need to know if it is in the past or the future; we can buy or sell an object therefore the money we use will be added or subtracted from the total; ...

5.17.3.1 Positive numbers

A positive number is a number different than zero, preceded by a "+" (plus) sign. Sometimes positive numbers are not preceded by any sign.

If a number is not preceded by a sign it is considered to be a positive number.

- Example of positive numbers: 5; +3; +7; ...

5.17.3.2 Negative numbers

A negative number is a number different than zero, preceded by a "–" (minus) sign. Negative numbers are always preceded by a "–" sign.

5.17.3.3 Absolute value of a number

The absolute value of a number is the value of the number without a sign. The absolute value of a number is written as shown:

- |a| is the absolute value of the number a and has a positive value.
- Example of absolute value of a number: |+9| is equal with 9; |-7| is equal with 7; |0| is equal with 0;.....

5.17.3.4 Fraction (ratio)

A fraction differentiates (or "measures") parts versus the whole.

- XX/YY
- XX is the part
- YY is the whole

In the English language, this read as: XX parts of YY. Note that the number representing the whole comes at the end.

In the Chinese languages, parts of a whole are stated as: YY 分之 XX. Note that the number representing the whole comes at the beginning. When expressing a fraction in Chinese, the whole (denominator) is always said before the part (numerator).

For example, in concern to the fraction 2/3:

- 2 is the part
- 3 is the whole

In the English language, it is read as:

- Two thirds
- Two [parts] of three [parts]
- Two over three
- Two [out] of three

In the Chinese languages, it is read as:

- From three [pieces] there are two [pieces]
- Three pieces, two

5.17.4 Percentages

In the Chinese languages, the same basic "part of whole" construction is used. In this case, the whole is 100 or 百 and the part is the actual percentage (as represented by an integer). Hence, the expression is as follows:

- 百分之 XX = %, where XX parts of one hundred (XX%)
- For example, 20% may be expressed as: from 100 there is 20%

5.17.5 Decimals

Fractions can be stated as decimals easily in Chinese. This is because each digit of the decimal fraction is stated individually. So instead of remembering tens, hundredths, thousands, etc., the numbers following the decimal point (to the right) are numerically listed.

The decimal point in Mandarin decimal fractions is stated as 点 (diặn). If the number begins with the decimal point, it can optionally be prefaced with zero or 令 (líng).

Here are two examples:

- 1.3= 一点三 (yī diặn sān)
- 0.5674= 零点五六七四 (ling diặn wŭ liù qī sì)

The West reads and writes from left to right but the numbers come from India and Arabia where one reads and writes from right to left. In English 14 is pronounced as fourteen but it should rather be ten·four. 21 is pronounced in proper order as twenty·one, but is better pronounced as two·ten·one, so that the decimal positional system is also supported by pronunciation.

5.17.6 Counting (value incrementing)

Counting from 0 to [1 through] 10 in English and Chinese generates ten unique words, which symbolize the ten numbers of the most widely used base-10 digit system (0-9). From 11–20, the English and Chinese ways of linguistically expressing count begin to differ.

In the English language, to count from 11–20, ten additional words are required. Hence, to count from 1–20 in English, 20 unique words need to be learned. In the Chinese language, to count from 11–20, no new words are introduced. Instead, the Chinese language reincorporates the same words used for 1–10, to cover all the numbers from 11–20. If "you" can count from 1–10 in Chinese," you" can count to 20 by default.

In the English language, to count from 21–100, eight new words are introduced (thirty, forty, fifty...,hundred.). In the Chinese language, to count from 21–100, only one new word is introduced: hundred. No new words are introduced to count from 11–99 in Chinese.

Therefore, to count from 1–100 in English, someone needs to account for 28 words. To count from 1–100 in Chinese, someone need only account for 11 words. This is a significant difference and impacts learning.

After a child learns to count from 1–10 in Chinese, one additional logical concept/rule is applied (iteration), and they can seamlessly count from 11–99. The logic stream remains, and the child doesn't have to learn a single new word to count from 11–99. By learning to count from 1–10, they have learned everything they need to count from 1–99. Further, there are no additional spelling complexities and exceptions that need to be learned.

Watch a child learn to count in English. What happens after they learn 1–10? They get confused, because it's ten new words to count from 11–20. And what happens after they count to 20? They often get stuck at each ten segment for the simple reason that it's a new word – thirty, forty, fifty, etc.

5.17.7 Measurement

How to express a certain amount or quantity [of something] is different in different linguistic expressions of numbers.

Counting in Chinese requires the use of a special class of words called "measure words". Some words in English perform similar functions, but the difference is that in Mandarin all words require a measure word when being counted. These serve to give units for counting and classifying nouns. In Chinese, there are 'measure words' for [almost] everything. Many of the measure words may be applied to multiple different [types of] objects. Every object has a measure word that must be known (and is to be used when expressing a measure concerning that object).

When counting objects, Chinese uses the following formula:

- Number + Measure Word + (Object)

When counting fractions of a thing, Chinese uses the following formula:

- Whole Number + Measure Word + Fractional Number + (Object)

For instance:

- For pens, it is said, "three sticks of pens".
- In English, one could say that there are three sheets of paper. In this case, the word "sheets" would act as a measure word. It would not make sense in Chinese to say, "three papers" (as in, "I have three pencils and three papers in my hand").
- It is not possible to say "one person" or "two people", as in English. Instead, it is said, one *ge* person or two *ge* people. Here, *gu/ge* is [one of] the measure word for person/people.

In English, counting things involves two inputs, the quantity, and the name of that which is being counted. For example,

- [There are] three pens.
- [There are] three pieces of paper.

English does, occasionally, use measure words.

- For instance, how many 'heads' of livestock are there? Herein, 'heads' is a measure word. There are nine heads of livestock; nine cows.

The linguistics around measurement in Chinese make quantification and comparison unnecessarily more complex by requiring users to remember and apply an additional word. Therein, that word may or may not be inherently relevant to the conceptual characteristics of the measured object itself. And, it certainly introduces the likelihood for conceptual confusion. For instance, in Chinese, the concept 'stick' is used as the measurement word for quantifying pens; as in, "three sticks of pens". To a large extent measure words pertain to an objects shape or a significant characteristic, but this is not always the case. For instance, there is a measure word pertaining to long flexible/flowy things, "tiao". For example, a fish, dragon, and pair of pants all use tiao as their measure word. There is a measure word for things that are flat, and this word is used for pieces of paper and for tables (even though tables are not flat, but have legs).

5.17.8 Statistics

Statistical methods are basically instruments for processing information. The information that statistical methods process is numerical in nature and derives from one or another of several forms of measurement. The various statistical processes/procedures make different assumptions about the incoming information (as in, measures) they are to process. If the available information incoming for processing is sufficient for the level of statistical operation selected, then the resultant output of that process will provide a high basis for forming rational conclusions. If the statistical operation processes information that violates its input criteria, then the result will be nonsense, and no rational conclusions can be formed.

> **APHORISM:** *Nonsense can still look elegant. Garbage in, garbage out. These are two saying from the language of computer engineers and programmers, an information processing system is only as useful as the accuracy and precision of the information it is fed for processing. A functioning system fed (input) wrong information will output wrong information.*

6 Measurement types

6.1 Biological organism taxanomical hierarchy

In biological classification, rank is the relative level of a group of organisms (a taxon) in a taxonomical hierarchy.

- Kingdom
- Sub-kingdom
- Infra-kingdom
- Division
- Subdivision
- Infra-division
- Class
- Order
- Family
- Genus
- Species

6.2 Measurement unit: illumination (rate of induction, light)

A prism breaks light into different wavelengths (wave[lengths]) of light.

6.2.1 Units of capacity

Units of capacity include, for example:

- Ounces
- Cups
- Pints
- Quarts

6.2.2 Linear density units

Linear density units include, for example:

- Grams per meter (g/m)

6.2.3 Measurement Unit: Time

No content here yet.

6.2.4 Time scales

1. Nominal time scale (nominal time of day) – AM, PM
   - Categories and no additional information.
2. Ordinal time scale (ordinal time of day) – morning, noon, afternoon, evening, night
   - Indicates direction or order of occurrence; spacing between is uneven.
3. Interval time scale (interval time of day) –12,1,2,3,4, 5,6,7,8,9,10,11,12,1,2,3,4,5,6,7,8,9,10,11,12
   - Equal intervals; difference between 1 and 2 pm is same as difference between 11 and 12 am.
4. Ratio time scale (ratio time of day) - 0,1,2,3,4,5,6,7,8, 9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24
   - 24-hour time has an absolute 0 (midnight); 14 o'clock is twice as long from midnight as 7 o'clock.

6.2.5 Solar time

Timekeeping. Why is a calendar the way it is? Why does the "year" start on January first? //The year of which belief system?

1. Currently, months don't follow the actual lunar cycles. Why are there 24 hours in a day.
2. From ~2:00 onward where the watch is shown he describes all possible time data.

6.2.6 Time and frequency metrology

This area of metrology studies components and their characteristics, especially

- Frequency standards
- Synthesizers
- Oscillators
- Digital clocks

6.2.7 Measurement Unit: Time

A clock is an instrument to indicate, keep, and co-ordinate time. The word clock is derived (via Dutch, Northern French, and Medieval Latin) from the Celtic words clagan and clocca meaning "bell". A silent instrument missing such a striking mechanism has traditionally been known as a timepiece.[1] In general usage today a "clock" refers to any device for measuring and displaying the time.

The timekeeping element in every modern clock is a harmonic oscillator, a physical object (resonator) that vibrates or oscillates repetitively at a precisely constant frequency. This object can be a pendulum, a tuning fork, a quartz crystal, or the vibration of electrons in atoms as they emit microwaves.

Analog clocks usually indicate time using angles. Digital clocks display a numeric representation of time. Two numeric display formats are commonly used on digital clocks: 24-hour notation and 12-hour notation.

- Years measured in decades (as 10).

6.2.8 Time measurement system

Time sequence – sequence of data ordered in time; an ordered sequence of data in the time domain. Data that are not time series are usually called cross-section. A time series of cross-sectional data is called panel data – for example, studying weight or income of a particular group ['cohort'] over time is panel data. There are three common graphical representations of time sequenced data:

- Step-wise constant – each step represents a new value. For example, the cost of some item over time.
- Discrete – For example, the number of items sold per day.
- Continuous – For example, a temperature measurement over time.

6.2.9 The daily cycle unit [time]: Clock

In early 21st century society there are two primary unit [time] clocks:

1. The Midday centric clock – This clock is divided into two 12 hour segments. The first 12 hours of the day are signified by “a.m.”, which is the acronym for, “anti meridian”, which is Latin for, “before midday. The second 12 hours of the day are signified by “p.m.”, which is the acronym for “post meridian”, which is Latin for “after midday”. The two segments are as follows:
    - 12:00am –11:59am is before midday
    - 12:00pm-11:59pm is after midday
    - The 24-hour clock – This clock is known by many names including: Under the 24-hour clock system, the day begins at midnight, 00:00, and the last minute of the day begins at 23:59 and ends at 24:00, which is identical to 00:00 of the following day. 12:00 can only be mid-day. This is the clock unit system universally used the current planetary transportation and logistics system (which has several names, including: Zulu time; Greenwich Mean Time (GMT); and Universal Standard Time).
2. 00:00 and 24:00 are midnight
3. 12:00 is midday
4. 23:59:59 is 1 millisecond before midnight
5. 00:00:01 is 1 millisecond after midnight

Given what is known, the 24-hour clock is more logical and more intuitive than a clock with two specific segments, which adds an additional unit of measure. The 24-hour clock only has the unit ‘Time’. In order to have a complete comprehension of a value given by the 12-hour am/pm clock, two units must be given: the time unit, and the a/p m (after/before midday).

> **INSIGHT:** *Time is awareness of change. Time is directionality [of the experience of existence]. It is the unidirectional vector of experience [as consciousness], also idiomatically known as, “the arrow of time”. Time is molecular decay.*

6.2.10 Measurement unit: Human anatomy

Human locomotion - Land area coverage: Average adult human walks about 1 kilometer in about 15 minutes (60m x 15...7

6.2.11 Measurement unit: Electromagnetics

You may not be able to see it, but this antenna is putting out light (Read: electromagnetic radiation). The reason an antenna is used for radio waves (light) and a filament for visible light is the relative wave-length of the two waves. They are both electro-magnetic radiation (or induction). Radio typically has a frequency between 100kHz and 100GHz, and hence, it has wave lengths between 3mm and 3km. Visible light (ROYGBV) in the green is about a 500nm (or 0.0005mm) wavelength. Visible light can be radiated with an antenna if it is made 250nm long, but this is only recently possible with nanosonic fabrication.

6.2.12 Measurement unit: Kelvin

Kelvin is temperature, nm is intensity. Then we have spectrum and wavelengths, which are also related to K and nm. The surface temperature of the sun is about 5780 degrees Kelvin, so the peak intensity of solar radiation is about 501 nanometers, which corresponds to the blue-green region of the spectrum. The sun’s actual color is white because the range of wavelengths it emits is broad. The sun’s light appears yellow to us, however, because of the way Earth’s atmosphere scatters light.

Correlated Color Temperature. CCT of a given light source characterizes the temperature of an absolutely black body that would radiate a similar spectrum. The hotter the black body, the higher will be the CCT and the more blue or “cold” will be the light. As an illustration, sunlight has a yellow tint, whereas blue giants - huge stars with high temperature of the surface: 10000K and above (Sirius, for example) - seem bluish even to the naked eye.

6.2.12.1 The Color Rendition Index (CRI)

Unfortunately this term is often interpreted wrongly. It characterizes the influence of light source on the perception of an object’s color. This parameter shows how correctly a light source with a particular CCT will deliver the color of an illuminated object, compared with an ideal source - an absolutely black body with the same color temperature. To determine the CRI, a set of 8 standard color samples is illuminated with the source and with the light of a back body with the same color temperature. If none of the samples change their color, CRI is equal to 100. The index reduces in inverse proportion to the number of color changes in samples. It is usually believed that a CRI above 80 is good. It is important to know, however, that CRI is calculated for light sources with a particular color temperature. It is not appropriate to compare a 2700K, 82 CRI light source with a 5000K, 85 CRI source.

Also note that CCT and CRI are only defined for full-spectrum light sources. The CRI of monochromatic light is close to zero, and its CCT cannot be calculated. Look at Fig. 15, Fig. 16 - you can see a wide spectrum, starting near 120nm and finishing around 3000nm. In this whole range a clear maximum is present, and most of energy is radiated in a narrow band of wavelengths. Radiation spectrum of a black body can never have the shape of a narrow-band spike, similar to the spectrum of a monochromatic light source, and therefore, calculation of CCT for such sources makes no sense.

The International Commission on Illumination (CIE) suggested a categorization of infrared radiation into the following bands: a) Near-Infrared, NIR / IR-A (700 nm to 1,400 nm or 0.7 µm to 1.4 µm; 215 THz to 430 THz); b) Short-Wavelength Infrared, SWIR / IR-B (1,400 nm to 3,00 nm or 1.4 µm to 3 µm; 100 THz to 215 THz); and c) Mid-Wavelength to Long-Wavelength MWIR / LWIR / IR-C (3,000 nm to 1mm or 3 µm to 1,000 µm; 300 GHz to 100 THz).

1. Kelvin temperature scale
2. Rankine temperature scale
3. Celsius temperature scale

6.2.13 Units of measurement for radiation

There are 4 [different] types of [physical] measurement for radiation:

1. Exposure
2. Absorbed dose
3. Dose equivalent
4. Radioactivity

There are 4 [different] types of [physical] measurement for transportation:

1. Distance
2. Speed
3. Movement source
4. Environmental geometry

7 Applicable measurement constants and equations

The seven defined constants are:

- The frequency of the ground-state hyperfine splitting of the caesium-133 atom - $\Delta\nu(^{133}Cs)_{hfs}$.
- The speed of light in vacuum - c.
- The Planck constant - h.
- The elementary charge - e.
- The Boltzmann constant - k.
- The Avogadro constant - N_A.
- The luminous efficacy - K_{cd}.

The seven defined SI units of measurement are:

- s - Second (time)
- M - Meter (distance)
- Kg - Kilogram (mass)
- A - Ampere; formerly known as 'intensity' (I) (electric current)
- K - Kelvin (temperature)
- mol - Mole (amount of substance)
- cd - Candela (intensity of light)

7.1 Units and relationships

Pressure is defined as force/area which is the same as momentum/area/time since F=dp/dt. Momentum flow would be the momentum passing through a unit area per unit time so it's the same units.

- The physical units for heat are Watts (W), Joules/second (J/s) or calories/second (cal/s)
- Heat is measured in watts.
- Heat flow is designated by the symbol q (Watts/m2).
- Electrical power is measured in watts.
- Power = work/time = J/s
- Energy is in joules.
- Electrical power is watts.
- Heat is not energy, but power.

Energy transfers are denoted by:

- Q = Transfer by Heat (J)
- W = Transfer by Work (J)
- q = Specific Transfer by Heat (J/kg)
- w = Specific Transfer by Work (J/kg)
- J = Transfer by Heat per Second, or Power (J/s = Watts)
- $\dot{W}$ = Transfer by Work per Second, or Power (J/s = Watts)

Internal energies and enthalpies are denoted by:

- U = Internal Energy (J)

- u = Specific Internal Energy (J/kg)
- H = U + PV = Enthalpy (J/K)
- h = u + Pv = Specific Enthalpy (J/kg.K)
- ṁ = mass flow rate (kg/s)

Single phase:

- Mechanically, power is calculated as leg pressure (Foot Pounds) times speed (Rotating Speed).
- Electrically, power is calculated as leg force (Voltage) times flow (Current).

Dual phase:

- Mechanically, power is calculated as leg pressure (Foot Pounds) times speed (Rotating Speed).
- Electrically, power is calculated as leg force (Voltage) times flow (Current).

Three phase:

- Mechanically, I'm not sure how to calculate the power.
- Electrically, power is calculated as cylinder force (Voltage) times flow (Current) times 1.732 (Square Root of 3).

- Coulomb = amount of electricity
- Coulomb = Ampere x Second
- 1 Coulomb = 1 Ampere • 1 Second
- 1C = 1A • 1s
- C = A • s

And,

- Coulomb = Farad x volt
- Coulomb = 1 Farad • 1 Volt
- 1C = 1F • 1V

7.2 Power

The unit of power is joules per second or J/s when work is measured in joules and time in seconds. The basic unit of power, 1 J/s is called a watt (W), named after James Watt who made important improvements to the steam engine. By definition, a watt is the consumption of one joule of energy per second. 1 W = 1 J/s

- Watts are units: units of power.
- A Watt is the unit of power.
- A Watt can be broken down further to the fundamental units of time, distance and mass.
- A Watt is 1 kg-m^2/s^3 in base SI units
- the power unit is 1 newton-metre/second, or 1 joule/second, this is 1 watt.
- A joule is a unit of work also known as force acting over a distance, i.e., F •d.
- Force is mass times acceleration, i.e., m•a.
- And acceleration is an exponential increase in distance over time, i.e., d/s^2.
- Watt = J/s = F•d/s = m•a•d/s = m•(d/s^2)•d/s = m • d^2 / s^3

- In metric that is kilogram•meter^2/second^3
- Or, generically, mass•distance^2/time^3 or ML^2/T^3
- A watt, as originally defined is volt²/ohm, the current dimensions are V²/R.
- If you use density-velocity-time, a watt = dv^5t^2.

Work Transfer:

- Work (J) = force (N) x distance (m)
- 1 Joule = 1 newton x 1 meter
- In units energy = joules (J)
- Work = joules (J)

Work is not a vector, but force and displacement are vectors.

- W = $^{+}$F x $^{+}$d - when force causes a displacement, work (energy) is positive (F x d = work)
- W = $^{+}$F x $^{-}$d - when force hinders a displacement, work (energy) is negative (F x $^{-}$d = $^{-}$work).
- W = $^{+}$F x ^{0}d - when force results in no displacement, there is no work (F x 0 = 0 work).

Secondly, work (W) is accomplished by a force (f) acting through a distance (d).

$W = \int f_i \cdot dx_i$ *(i.e., Work = Force x Distance)*

For a constant force; the work done formula as force x distance, only applies if you have a constant force:

$W = f_i \cdot \Delta x_i$

NOTE: *Mass is simply how much stuff there is in the object. No matter where you put an object in the universe without taking it apart or breaking it, the mass will always be the same. However, the weight changes. Weight is relative to the field in which the mass exists.*

7.3 Units of energy and power

In order to predict and account for "action", energy is a required quantification. In physics, action is an attribute of the dynamics of a physical/material system. Action is understood as a mathematical functional that takes the trajectory, also called path or history (memory), of the system as its argument and has a real number as its result. Generally, the action takes different values for different paths. Action has the dimensions of [energy]·[time\memory], and its SI unit is joule-second. This is the same unit as that of angular momentum.

Energy and power are measured in a variety of ways depending on the system (and scale) in which the measurement is occurring.

- Energy determined to be contained in a system is called static form of energy (e.g. internal, kinetic, potential energies).
- Dynamic forms of energy come from energy interactions, where energy crosses the system boundary during a process (e.g. heat transfer and work).

Electron volt (eV) is a unit of energy, not voltage. The amount of energy expressed when an electron is accelerated through a potential of 1 volt.

- e = charge on the electron = 1.6×10^{-19}C
- 1V = 1J/C
- eV = (1.6×10^{-19}C) x (1J/C) = 1.6×10^{-19} J

A measure of energy can be expressed/signified in the following ways (i.e., the direct release of energy is measured in units of):

- Electron-volt (eV) - A unit of energy equal to the work done on a charge ("electron") in moving it through a potential difference of one volt. An electron volt is defined as a unit of energy. An electron volt is the energy an electron gains when it is accelerated through a potential difference of one volt. Electron-volt scales: Nuclear energy scales are MeV; Chemical energy scales are eV.
- Joule or jule (J) = a unit of work (energy) equal to the work done when the point of application of a force of one Newton moves a distance of one meter, in the direction of the force. One joule is defined "mechanically" as the energy transferred to an object by the mechanical work of moving it a distance of 1 metre against a force of 1 newton (i.e., netwon-meter).
    - $1\ J = 1\ ((kg \cdot m^2) / s^2)$
- Watt-seconds or Watt-hour (KWh)
- Calorie
- Radiant energy units
- Heat units (e.g., British thermal units, BTUs)
- Electromagnetic energy units (SI electromagnetic units)
- Nuclear energy units
- Energy - The ability or potential to do work.
- Work - The transfer of energy from one carrier to another.
- No movement = no work.
- Power - The rate at which work is done and energy is transferred.

> **NOTE:** *A joule is a rather small amount of energy, roughly equal to the kinetic energy of a very gently tossed baseball, or to the gravitational energy that you give to a baseball when you lift it by 70 centimeters.*

The more Kilowatts used, the more energy that's being used up.

A kilowatt is 1,000 watts; one watt is the same as one Joule per second (J/s). Which is confusing, since J/s mentions a time frame (second) but it doesn't compare to kWh (which mentions hours, but isn't about time).

Watts cannot be converted to amps, because watts are power and amps are coulombs per second.

If you have at least two of the following, then the missing one can be calculated: amps, volts, watts.

- Watts = amps x volts
- Current = wattage / voltage
- Voltage = wattage / current

Amps are how many electrons flow past a certain point per second. It is equal to one coulomb of charge per second, or 6.24 x 10^18 electrons per second. Volts is a measure of how much force that each electron is under, which we call "potential". Power (watts) is volts times amps. A few electrons under a lot of potential can supply a lot of power, or a lot of electrons at a low potential can supply the same power.

7.4 Energy and work relationship

Energy is substance-like, and work is a transfer mode of that substance. However, energy and work are the same unit of measure although they are not necessarily measuring the same thing.

- Linear kinetics
    - Work = Δtotal mechanical energy
    - Assuming a rigid body that cannot store elastic energy:
    - $Fd = \Delta(.5mv^2 + mgh)$
    - $Fd = \Delta.5mv^2 + \Delta mgh$
- F = ma
    - Work = $m(.5v^2)$
- Angular kinetics
    - Work = Δenergy
    - $Fd = \Delta.5mv^2$
    - $\tau\Theta = \Delta.5I\omega^2$

A watt is a watt is a watt whether it's electrical or mechanical or chemical.

> **QUESTION:** *How much energy is the something (e.g., a bulb) using? That depends on time -- how long it is operating.*

7.5 Kinetic energy systems

- Energy = .5 mass (m) · velocity $(v)^2$

7.6 Potential energy

- Potential energy is often thought of as "stored" kinetic energy, meaning that bodies remain stationary in a potential field while held in place by some force, and upon change in this force (such as breaking the twig holding an apple, or breaking the bond between two atoms), potential energy is converted to kinetic form (the apple "falls" or the molecule "dissociates").
- Gravitational potential energy
  - Energy = mass (m) · gravity (g) · height (h)
- Units: Joules

Gravitational potential energy - energy contained in an object due to its vertical position above the plane of the Earth.

- Gravitational potential energy (PE) = m x g x h
- Where m=mass, g = gravitational constant 9.8m/s^2, h=height
- g is known as the gravitational constant. It measures the strength of the Earth's gravitational pull on falling objects. Falling objects accelerate downwards at a rate of 9.8m/s^2

Gravitational potential - potential energy per mass.

- PE/mass = (m x g x h) / m

Gravity analogy:

1. Two points in space at the same height\coordinate have zero potential difference.
2. 2. $\Delta\Phi := \Phi(x_2,t0) - \Phi(x_1,t0)$
   A. A falling rock - A 1 kilogram rock (unit mass) can transfer more gravitational potential energy to kinetic energy if it "falls" off the side of a 100 metre ledge than if it falls off a 10 metre ledge.
   B. A falling electron - Similarly, a 1 coulomb charge (unit charge) can transfers more energy if it "falls" through an electrical potential difference of 100 volts than if it falls through 10 volts.
   C. The rock "falls" through a gravitational potential difference and the coulomb "falls" through an electrical potential difference.

7.7 Total energy

Total energy, E. Energy can generally be divided into two groups:

- Macroscopic: energy a system possesses as a whole with respect to some outside reference frame., such as kinetic energy (KE) and potential energy (PE).
- Microscopic: energy related to the molecular structure of a system and the degree of the molecular activity, and they are independent of an outside reference frame. The sum of all forms of microscopic energy is called the internal energy, U, of the system.

The **internal energy** of a system is comprised of:

- Sensible energy: the portion of internal energy associated with the kinetic energy of molecules (i.e. translational, rotational, and vibrational kinetic energies).
- Latent energy: intermolecular forces between the molecules of a system.
- Chemical (or bond) energy: internal energy associated with atomic bonds in a molecule. During combustion processes, atomic bonds are broken and new ones are formed, altering the internal energy of the system.
- Nuclear energy: energy harnessed from the bonds within the nucleus of an atom.

Mechanical work is defined by the relation w = Fdx, where w = work is done, F is force, x is displacement, and the subscripts i and f denote the initial and final states respectively. Similarly, mechanical power is defined as P = Fdv where P is power delivered and v is velocity. Barring, special energy considerations (e.g. magnetic, chemical, surface-tension, etc ...), the total energy of a system can be expressed as:

- E = U + KE + PE

7.8 Internal energy

Internal Energy (E) measures the energy state of a system as it undergoes chemical and/or physical processes. Like other thermodynamic variables, internal energy exhibits two important properties:

1. It is a state function, and
2. it scales as an extensive.

Being a state function means that E has the following property:

- E = Ef - Ei

The relationship between the internal energy of a system and its heat and work exchange with the surroundings is:

- E = q + w

7.8.1 Energy

- Energy = force x distance
- Force = pressure x area

- Distance = volume / area
- Energy = pressure x volume (psig x cu-in = in-lbs)

7.8.2 Work

Mechanical work is:

- Work is scalar.
- Work is Joules.
- Mechanical work is force through a distance (displacement):
  - W is the work done, F is the force, d is the displacement, and · indicates the dot product.
  - Work (W) = Force (F) · distance (x)
  - W = ∫F • x
  - W = F x d x cosΘ
  - Units: Joules (do not use N.m)
  - Force (newton) = mass x acceleration
  - A force has both magnitude and direction, making it a vector quantity. It is measured in the SI unit of newtons.
- The unit of power is the joule per second (J/s), known as the watt, named confusingly after James Watt.
- The mechanical shaft power P in Watts applied to a generator is given by:
  - P = ωT
  - Wherein, ω is the speed in radians per second and T is the torque in Newton meters.
  - Therein,
  - Work = torque x revolutions

7.8.3 Mechanical energy

Mechanical energy is:

- Mechanical energy = Kinetic energy + Potential energy
- Mechanical energy = $1/2mv^2 + mgh$

7.8.4 Mechanical pressure

Mechanical pressure is:

- Pressure is defined as the normal force exerted by a fluid per unit area. Pressure can exist, even if no work is being done -- pressure has units of N/m^2 (Pascal as newtons per meter squared).
- The basic unit of mechanical pressure is the newton per square metre.
- Pressure (P) = force/area
- 1 Pa = $1N/m^2$

The pascal (symbol: Pa) is the SI derived unit of pressure used to quantify internal pressure, stress, Young's modulus and ultimate tensile strength. It is defined as one newton per square metre.

NOTE: *The flow of momentum is pressure.*

7.8.5 Mechanical power

- Mechanical power (P) is the quotient of mechanical work (A) by time (t).
- P = mechanical work (A) / time (t) = (F x s) / t
- The SI unit of measurement is watts (W).
- Rotational mechanical power is torque (T) and angular velocity (ω) (see Rotational speed).
- P = Tω

7.8.6 Mechanical and fluid systems

In these systems, 'work' is another word for 'energy'. Work (W) = Force (F) · distance (x)

- Unit of force = Newton
- 1 Newton accelerates a mass of 1 kg. by 1 m/s2 (in case of no friction).
- A mass of 1kg on earth experiences a gravitational force of 9,8 Newton.
- Unit of work (energy) = Joule
- 1 Newton moves an object over 1 meter, the required amount of energy is 1 Joule.

NOTE: *Torque is force at a distance. Torque is a pseudovector (equivalent to a mathematical bivector in three dimension); energy is not. A pseudovector is distinguished it from a true polar vector. The units for torque are Newton-meters. Although this is algebraically the same units as Joules, Joules are generally not appropriate units for torque. Torque is usually given by rFsinθ, not just rF, unless the angle is always 90˚of course because sin90=1.*

7.9 Electrical systems

Energy is a quantity indicating the capacity to do work.

- Energy = Power x Time

Power is the rate at which work is done.

- Power = Energy / Time

Voltage exists if charges are moving [through] a distance. Voltage is electric [potential] energy per unit charge, measured in joules per coulomb (= volts).

- Voltage (V) = *energy* in joules (J) / *charge* in coulombs (C)
- Voltage (V) = joule (J) / coulomb (C)
- 1V = 1J/C
- Potential = the *ability* to do work.
- Electric potential is the ability to do [electrical] work per electric unit.
- Similarly, the electric field is electric force per

charge. E = f/q \ f=qE

Electric potential energy - work required to move a charge.

- $F = (k\ 0_1 0_2)/d^2$

Electric potential - The electric potential Φ refers to a quantity with some numeric value. Expresses the effect of an electric field of source in terms of the location within the electric field.

- Φ = PE / q

Electric potential difference (ΔV) - the difference in electric potential (V) between the final and inital location when work is done upon a charge to change its PE.

- $\Delta V = V_B - V_A$ = work/charge = ΔPE/charge

The analog is:

- Volts (V) = Height or head (H)
- Charge (q) = mass m
- Current I = Δq/Δt = rate of mass flow Δm/Δt
- Power = VI = ghΔm/Δt
- Energy = VIdt = Vq = ghm // Energy is the time integral/sum of power.

7.9.1 Electrical work

Electrical work is:

- We = VI
- V = Voltage
- I = current
- We = VIΔt
- t = time
- Power = Energy/Time
- Energy = Power × Time
- Energy (J) = volts · charge in coulombs
- Power (w) = volts · amps
- The standard unit of electrical power is the Watt, which is defined as an [electric] current of one ampere, pushed by a voltage of one volt.
- Watt or kilowatt (watt/1000)
- Current (I) = charge in coulombs / time in seconds
- 1 W = 1J/s
- 1 kW = 1000 W = 1000 J/s
- 1 MW = 1,000,000 W = 1,000,000 J/s

Electric current is measured in coulombs per second (amperes or amps; A).

- 1 Ampere is equal to 1 Coulombs per second.

Current is rate of change in the electric field:

- current (I) = Δq/Δt
- wherein, q=charge

7.9.2 Volt

The volt is defined as the energy transfer per coulomb of charge as charges move between two points in a circuit.

- V = ΔW / ΔQ
- i.e. energy change per unit charge (so that 1 V = 1 J C-1)

7.9.3 Power

Power is equated in multiple ways:

- Power = energy / time (Units: Watts (J/s))
- Power = pressure x volume/time
- Power = Δwork / Δtime
- Power = (force x Δdistance) / Δtime
- Power = force x velocity
- power = energy/time
- power = work/time
- power = (force x distance)/time
  - distance/time = speed
- power = force x (distance/time)
- power = force x speed

7.9.4 Fluid power system

Power in fluid systems is equated in multiple ways:

- Force (F) = pressure (P) · area (A)
- Pressure (P) = Force (F) / area (A)
- Fluid pressure (P) = force (F) / unit area (A)
- Fluid flow rate (Q) = volume (V) / unit time (A)
- Fluid power = pressure (P) x flow rate (Q)

7.9.5 Fuel systems

Fuel systems: In these systems, energy density is key.

7.9.6 Battery systems

No content here yet.

7.9.7 Pressure system

- Energy = pressure · volume
- Pressure = Force/ area unit

7.10 Unit conversion factors

Data and measurements may be expressed in any units, usually chosen for convenience of size. But when this data is used in physical equations, it must be converted to the units required by the coherent system chosen. Units must also be converted when translating from one coherent system to a different one.

Unit conversions begin with equations which relate sizes of units, for example: 1 meter = 3.28 feet. This equation states that the measurement "1 meter" is equal (equivalent to) the measurement "3.28 feet." To write simply 1=3 would be incorrect.

Equations relating measurements are manipulated by the ordinary rules of algebra, and the units are carried along according to the same rules. For example, if both sides of Eq. (3) are divided by 1 yard, the result is:

- 1 = 3 feet / 1 yard = 3 feet/yard
- 1 = 3 feet/yard

This last expression represents an identity relation for measurements. It is called a 'conversion factor'. In algebra it is often convenient to multiply an expression by another expression which is equal to one. When doing unit conversions, expressions may be multiplied by conversion factors, since they are physically equal to one.

Conversion factors for energy units:

- 1 kWh = 3,413 Btu
- 1 kWh = 3,600,000 joules
- 1 joule = 1 watt-second
- 1 joule = 1 Newton-meter
- 1 Btu = 1,055 joules
- 1 Therm = 100,000 Btu = 29.3 kWh
- 1 calorie = 4.184 joules
- 1 Btu = 252 calories

Conversion factors for power units:

- 1 watt = 1 joule/second
- 1 watt = 3.413 Btu/h
- 1 Btu/h = 0.2931 watt
- 1 kW = 1,000 watts
- 1 megawatt (MW) = 1,000,000 watts
- 1 kW = 3,413 Btu/h
- 1 ton of cooling = 12,000 Btu/h
- 1 horsepower (electric) = 746 watts

Guide for common fuels:

- Natural gas: 1,000 Btu/cu. ft.
- Propane: Between 91,333 Btu/gallon and 93,000 Btu/gallon
- Fuel oil: Between 138,700 Btu/gallon and 140,000 Btu/gallon
- Kerosene: Between 120,000 Btu/gallon and 135,000 Btu/gallon
- Gasoline: Between 114,000 Btu/gallon and 125,000 Btu/gallon
- Coal: 25,000,000 Btu/ton
- Seasoned dense hardwood firewood: Between 21 and 26 million Btu/cord
- Seasoned pine firewood: Between 14 and 16 million Btu/cord

Conversion factors used for measuring natural gas:

- 1 ccf ("centi- cubic feet") = 100 cubic feet
- 1 cubic foot of natural gas = 1,000 Btu = 0.01 Therm
- 1 Therm = 1 ccf of natural gas = 100,000 Btu = 29.3 kWh

Conversion factors for air pressure units:

- 1 atmosphere = 14.7 lb./sq. in. = 760 mm. of mercury = 406.78 in. of water = 101,325 Pascals
- 1 Pascal = 0.00401 in. of water
- 1 lb./sq. in. = 6,894.76 Pascals
- 1 lb./sq. ft. = 47.88 Pascals

> **NOTE:** *In the market, electrical energy is a measurable quantity that can be bought by the kilowatt-hour (KWh).*

> **NOTE:** *Energy density = electron-volt per cubic centimeter of space, or eV/cm*^{3}

> **NOTE:** *The basic quantity of electric charge is the electron. Conversely, electromagnetic waves have no charge.*

In the SI system of units, the joule (J) is a unit of energy, but the electron-volt (eV) is the traditional unit used in ion-solid interactions: 1 eV is defined as the kinetic energy gained by an electron accelerated through a potential difference of 1V. The electron-volt is a unit of energy. The definition of an electron volt is the kinetic energy a single electron acquires when moving through an electric potential of 1V. The charge on the electron is $1.602x10^{19}$ J. Commonly used multiples of the electron-volt are the kilo-electron-volt (10^3 eV) and the mega-electron-volt (10^6 eV).

Energy density units for problems involving thermodynamic analysis are typically in the form of joules per mole, where a mole (mol) represents Avogadro's number of particles or molecules: N_A = 6.02 x 10^{23} particles/mol.

Joule as a measure of energy. In particle physics, however, we use something more convenient called electron volt (eV) instead.

An electron-volt (eV) is the energy or work required to move an electron against a potential difference of one volt.

Scholarly references

- Garcia, F., Bertoa, M.F., et al. (2006). *Towards a consistent terminology for software measurement*. Information and Software Technology 48 (2006) 631–644. [biblioteca.uma.es]

- Schadow, G., McDonald, C. J., Suico, J. G., Föhring, U., & Tolxdorff, T. (1999). Units of measure in clinical information systems. Journal of the American Medical Informatics Association : JAMIA, 6(2), 151–162. doi:10.1136/jamia.1999.0060151 [www.ncbi.nlm.nih.gov]

Book references

- Colingnatus, T. (2018). *A child wants nice and no mean numbers*. ISBN 978-946367275-7 [thomascool.eu]
- Bently, J.P. (2005). *Principles of measurement systems*. Pearson: Prentice Hall. [research.iaun.ac.ir]
- Nicholas, J.V., White, D.R. (2001). *Traceable temperatures: An introduction to temperature, measurement, and calibration*. Wiley.

Online references

- Cheng, L. (2009). *How Language And Math Intersect: Chinese v. English*. Thinking About Thinking. [larrycheng.com]
- Johnson, J. *Schema.Org Measurement Proposal*. W3. Accessed: January 7, 2020. [w3.org]
- *Math Learning Tip: Find An Idea's True Name*. Better Explained. Accessed: January 7, 2020. [betterexplained.com]
- *Solar time*. (2019). Wikipedia. [wikipedia.org]
- Stewart, M. (2016). *Counting down to the new ampere*. Phys.org [phys.org]

TABLES

Table 43. Measurement > Quantity > Length: *Spatial length accounting for function.*

Space Used	Surface area required m^2/person	Number of levels	Project area m^2	Estimated height,m	Volume m^3
Dwelling					
Assembly					
Recreation					
Entertainment					
Storage					
Transportation					
Park					
Waste and water treatment and recycling					
Electrical supply and distription					
Cultivational areas					
Mechanical subysystems					
Communications system					
...					

Table 44. Measurement > Quantity: *Quantities per area unit.*

Quantities	Per area	units	
Density of magnetic induction	Φ/A	per cm^2	
Density of dielectric induction	(Q)/A	per cm^2	
Density of electrification	(x)/A2	per cm^4	
Formula	**Type**	**Unit label**	**Description**
x/T = W	Work or energy	Joule	Quantity of electrification varied with respect to time. In time its quantity changes and that is called work or energy. Note that energy is a derivative and does not have a primary existence.
Φ/T = E	Electromotive force	Volt	Total quantity of magnetism varied with respect to time. A volt is the rate at which magnetism is produced or consumed in an electrical system.
x/T = I,	Magnetomotive force	Ampere	Total quantity of dielectrification (produce or consume a dielectric field) and vary that with respect to time.
Q/T^2=P	Power or activity	Watt	Quantity of electricity (the product of Φ•x=Q) and vary it to the time squared.
Φ/I = L	Magnetic inductance	Henry	Magnetism compared to how much current required to produce it.
x/E = C	Dielectric capacity	Farad	For every quantity of dielectric field there has to be a certain amount of electromotive force that gives rise that field.
E/I = Z	Impedance	OHM	
I/E = Y	Admittance	Siemens	
L/T = R	Resistance, Henry per second	OHM	
C/T = G	Conductance, Farad per second	Siemens	
L • C = T^2 ($time^2$)	$^2\sqrt{LC} = T = F^{-1}$	$Hertz^{-1}$	Time rate of energy exchanged from the magnetic and dielectric field as they constantly dump one into another

TABLES

Table 45. Measurement > Electricity*: Electricity and magnetism physical units.*

Physical Units: Electricity And Magnetism						
Quantity and Definition	Electro-static (esu)	emu/esu	Electromagnetic emu	MKS/emu	Rationalized MKS	esu/MKS
Charge (Q)	statcoulomb	1/c	abcoulomb	10	coulomb	c/10
Current I = Q/t	statampere	1/c	abampere	10	ampere	c/10
Potential V = W/Q	statvolt	c	abvolt	10^{-8}	volt	10^{8}/c2
Resistance R=V/I	statohm	c^{2}	abohm	10^{-9}	ohm	10^{9}/c2
Capacitance C = Q/V	statfarad	$1/c^{2}$	abfarad	10^{-9}	farad	10^{6}/c
Electric field strength E=F/Q=V/s	dyne/statcoulomb = statvolt/cm	$1/c^{2}$	abvolt/cm	10^{-6}	volt/meter	10^{6}/c
Magnetic flux	erg/ statampere	c	maxwell	10^{-8}	weber = volt x sec	10^{8}/c
Magnetic induction	dyne/ (statamp x cm)	c	gauss	10^{-4}	weber/meter2	10^{4}/c
Magnetic field intensity	statampere/cm	1/c	oersted	10^{3}/4pi	ampere/meter	12pi10^{7}
Inductance	stathenry = statohm x cm	c^{2}	abhenry	10^{-9}	henry	109/c^{2}

Table 46. Measurement > Quantity *Sub-conceptualizations (as a classification scheme) of the concept, 'quantity'.*

Sub-concepts [for the concept 'quantity']		Sub-conceptual application
length, l	radius, r	radius of a circle A, r_A or r(A)
	wavelength, lambda	wavelength of the sodium D radiation, λ_D or λ(D; Na)
energ, E	kinetic energy, T	kinetic energy of particle i in a given system, T_i
	heat, Q	heat of vaporization of sample i of water, Q_i
electric charge, Q		electric charge of the proton, e
electric resistance, R		electric resistance of resistor i in a given circuit, R_i
amount-of-substance concetration of entity B, c_B		amount-of-substance concentration of ethanol in wine sample i, c_i(C2H5OH)
number concentration of entity B, C_B		number concentration of erythrocytes in blood sample i, C(Erys; B_i)
Rockwell C hardness (150 kg load), HRC(150 kg)		Rockwell C hardness of steel sample i, HRC_i (150 kg)

Table 47. Measurement > Units*: Energy and power in base formula.*

Type	Symbol	Description	In Water	In Electrical Energy	Base Units
Energy	E	The ability to do work	Power=Current*Pressure (P=Q*H)	Power=Current*Voltage (P=I*V)	kg•m^2/s^3
Power	P	Rate at which work is done	Energy=Power*Time (E=P*t)	Energy=Power*Time (E=P*t)	kg•m^2/s^2

TABLES

Table 48. Measurement > Metrological: *Metrological units.*

	Units						
Descriptive Elements	Second (s)	Kilogram (kg)	Candela (C)	Kelvin (K)	Ampere (A)	Meter (m)	Mole (mol)
Measures	**Time**	**Mass**	**Luminous intensity**	**Temperature**	**Current**	**Length**	**Amount of substance**
Requires / Based Upon	Hyperfine-transition frequency of the caesium-133 atom (ΔVCs)	Planck's constant (h)	Luminous efficacy of monochromatic light of frequency 540 x 10^{12}Hz and a radiant intensity of 1/683 watts per steradian (Kcd)	Boltzmann's constant (k)	Charge on the electron (e)	Speed of light in a vacuum (c)	Avogadro's constant (N_A)
Definitions / Constant Used	Duration of 9,192,631,770 cycles of the radiation corresponding to the transition between two hyperfine levels of caesium-133	One kilogram is Planck's constant divided by 6.626 070 15 x 10^{-34} m^{-2}s	Luminous intensity of a light source with frequency 540 x 10^{12} Hz and a radiant intensity of 1/683 watts per steradian	Equal to a change in thermal energy of 1.380 649 x 10^{-23} joules	Electric current corresponding to the flow of 1/(1.602 176 634 x 10-^{19}) elementary charges per second	Length of the path traveled by light in a vacuum in 1/299,792,458 seconds	Amount of substance of a system that contains 6.022 140 76 x 10^{23} specified elementary entities

Table 49. Measurement > Energy: *Common units of energy.*

Common Units Of Energy And Power	
Energy	Power
joule	joule/sec
calorie	calorie/min
Btu	Btu/hour
watt-hour	watt
kilowatt-hour	kilowatt
orange	orange/day

Table 50. Measurement > Motion: *Linear and rotational motion as speed and force.*

	Speed	Force
Linear motion	speed s	force f
Rotational motion	angular speed ω	twisting force τ

Table 51. Measurement > Units > Transfer: *Conserved quantities and rates of transfer.*

Conserved Quantity		Rate of Transfer	
Name	Units	Name	Units
energy	joules (J)	power	watts (W)
momentum		force	newtons(N)
angular momentum		torque	newton-meters

Table 52. Measurement > Units: *Linear and rotational work and power.*

System	Work	Power
Linear	W = F x d	P = W/t P = F x d/t = F x v
Rotational	W = T x Θ	P = W/t P = T x Θ = T x ω

TABLES

Table 53. Measurement > Units*: Generalized table of units of function.*

Description	Energy	Work	Force	Power	Pressure
Measured in units called			Force & torque are measured	Calculated	
Instrument of measurement is a			Dynamometer		Manometer
Has or does not have subcategories	Yes. Two primary forms (kinetic & potential). Multiple forms and types.	Manager...just joking.	Yes. Mechanical contact forces - normal, applied, friction, tension, spring, resisting. Electromagnetic force. Gravitational force. Nuclear force(s). Mechanical twisting force - torque. In mechanics, forces cause linear motion, torques cause rotational motion. Curved motion has centripetal and centrifugal force (and coriollis force).	Yes. Electric, mechanical, fluid, thermal.	No.
Formula(s)		Work = Force x Displacement	Force = Mass x Acceleration (Or) Force = dP/dt (change in momentum by time)	Power = work done/ time taken	
Definition	Measure of ability to do work. It doesn't mean work is being done, but that work can be done.	Change in energy via force. As a result of application of the force, if the configuration of the system changes, the measure of the same is the work done (force into displacement).	An influence that interacts to change the motion of an object. Cause of change in state of motion.	Rate of energy transfer by doing work. Power is the rate of doing work or expending energy. Rate of work done or the rate of energy release.	
Definition with respect to motion	Energy is the magnitude of stress, introduced in universal medium during work.	Work is the magnitude of distortions, introduced in universal medium about a 3D matter-body.	Force is matter-content times rate of change of work-done or rest mass times acceleration.	Power is temporal rate of work-done during acceleration.	
Value type	Scalar (given that work is scalar). Conserved.	Scalar (scalar but no direction)	Vector (direction) and magnitude	Since Energy and Time are both scalars, Power is a scalar also.	Scalar (magnitude and no direction)
Observable when?		When energy transfers.			
Linear motion		W = FΔ x or W = f x dx		P = Fv	
Rotational motion		W = tΔ0 or w = t x d0		P = tw	
Curved motion					

TABLES

Table 54. Measurement > Dimensionality: *Table shows electrical dimensions.*

Note here that there is disagreement over the naming of the electric field. Steinmetz eliminated the use of the term "electric field", and instead, called it the 'dielectric field'. The usage of the term 'magnetic field' is not in disagreement. In Steinmetz's electrical theory, electricity has to be the product of total magnetization times total dielectrification. If it is just one or the other it is not electricity. Hence, a charged capacitor with a total dielectrification and no total magnetization is not electricity. It is only when the energy of each is exchanged in a cyclic process that electricity is appears. In the days of Franklin, metals were called non-electrics because they destroyed the [di]electric field. Energy can be taken apart and put back together. Dimension - one of a group of properties whose number is necessary and sufficient to determine uniquely each element of a system of entities. The misuse of the word dimensions arrives when the term is defined as directional measurement or number of coordinates (i.e., 3 dimensional space). In reality, space is a single dimension (i.e., there exists, only the dimension of space). Space-time then is the relation of two distinct dimensions: the single dimension of space and the single dimension of time. For instance, velocity is expressed as the ratio of the dimension of space to the dimension of space (distance/time=velocity). Thus, velocity is expressed as a two dimensional relationship. Capacitance is a type of electrical energy storage in the form of field in an enclosed space. This space is typically bounded by two parallel metallic plates or two metallic foils on an intervening insulator or dielectric. A nearly infinite variety of more complex structures can exhibit capacity, as long as a difference in electric potential exists between various areas of the structure. The oscillating [tesla] coil represents one possibility as to a capacitor of more complex form, and will be presented here. All the lines magnetic force are closed upon themselves. All the lines of dielectric force terminate on conductors, but may form closed loops in electromagnetic radiation (EMR). Any line of force cannot just end in space. Inductance represents energy storage in space as a magnetic field. The lines of force orientate themselves in close loops surrounding the axis of current flow (magnetism scraping on the wire) that has given rise to them. The large the space between this current and its images or reflections, the more energy that can be stored in the resulting field. Inductance in electronics is electrical inertia. quantity dimensions vs. space and time as metrical dimensions.

Quantity in undivided form		
q	Total Electrification	Plank
Φ	Total Magnetization (outerspace aspect of dielectricity)	Weber
Ψ	Total dielectrification (innerspace aspect of dielectricity)	Coulomb
Basic relationship		
q/Ψ = Φ	Magnetic induction	Weber
Φ x Ψ = q	Magnetism and dielectricity are the two components of electricity	Plank
q/Φ = Ψ	Dielectric induction	Coulomb
Derivatives of quantity by space, A		
Φ/A =	Density of magnetic induction	per cm^2
Ψ/A =	Density of dielectric induction	per cm^2
q/A^2 =	Density of electrification	per cm^4
Derivatives of quantity by time, t		
q/t = W	Work or Energy *The quantity of electrification varied with respect to time. Energy does not have a primary existence. Energy is a derivative.*	Joule
Φ/t = E	Electromotive force *The quantity of magnetization (magnetic field) varied with respect to time. A 'volt' is the rate at which magnetism is produced or consumed in an electrical system.*	Volt
Ψ/t = I	Magnetomotive force *The quantity of dielectrification varied with respect to time (i.e., a dielectric field is either produced or consumed, and it is varied with respect to time.*	Ampere
q/t^2 = P	Power or/of Activity	Watt
Proportionality		
Φ/I = L	Magnetic inductance	Henry
Ψ/E = C	Dielectric capacity	Farad

TABLES

E/I = Z	Impedance	Ohm
I/E = Y	Admittance	Siemens
Density of decay		
L/T = R	Resistance - *The destruction of energy in an electrical system in Henrys per second.*	Ohm
C/T = G	Conductance - *The creation of energy in an electrical system in Farads per second.*	Siemens Mho
$LxC=t^2$	$^2\sqrt{LC} = t = F^{-1}$ *Frequency of oscillation (time rate between the two fields as they "dump" into one another.*	$Hertz^{-1}$

Table 56. Measurement > Units*: Fundamental (base) quantities, dimensions, and units.*

Dimension type	Name of physical quantity	Unit name	Symbol / Abbreviation
Temporal dimension	Time	Second	s
Linear dimension	Length	Meter (Metre)	m
Matter dimension	Mass	Gram (Gramme)	g
Electric dimension	Electric current	Ampere (formerly known as Intensity)	A
Thermodynamic dimension	Temperature	Kelvin	K
Atomic mass dimension	Atom[ic amount of substance]	Mole	mol
Inductive illumination dimension	Illumination	Candela	cd

Table 55. Measurement > Units*: The expression of kinematical units in terms of units of energy.*

Quantity	Dimension		Conversions
	SI Units	Natural Units	
Mass	Kg	E	1 GeV = 1.8×10^{-27} kg
Length	M	1/E	1 GeV-1 = 0.197×10^{-15} m
Time	S	1/E	1 GeV-1= 6.58×10^{-25} s
Energy	Kg m^2/s^2	E	1 GeV = 1.6×10^{-10} Joules
Momentum	kg x m/s	E	1 GeV = 5.39×10^{-19} kg x m/s
Velocity	m/s	None	1 = 2.998×10^{8} m/s (c)
Angular momentum	kg x m^2/s	None	1 = 1.06×10^{-34} J x s (ħ)
Cross-section	m^2	1/E2	1 GeV^{-2} = 0.389 mb = 0.389×10^{-31} m2
Force	kg x m/s^2	E2	1 GeV2 = 8.19×10^{5} Newton
Charge	C-As	none	charge C=A x s none 1 = 5.28×10^{-19} Coulomb; e=0.303=1.6×10^{-19} C

TABLES

Table 57. **Measurement > Units***: The most common SI derived units.*

[name] Derivation (derived quantity)	[label] Unit Name	Unit Symbol	Expression in terms of SI base units
dynamic viscosity	pascal second	Pa s	m^{-1} kg s^{-1}
moment of force	newton metre	N m	m^2 kg s^{-2}
surface tension	newton per metre	N/m	kg s^{-2}
heat flux density, irradiance	watt per square metre	W/m^2	kg s^{-3}
heat capacity, entropy	joule per kelvin	J/K	m^2 kg s^{-2} K^{-1}
specific heat capacity, specific entropy	joule per kilogram kelvin	J/(kg K)	m^2 s^{-2} K^{-1}
specific energy	joule per kilogram	J/kg	m^2 s^{-2}
thermal conductivity	watt per metre kelvin	W/(m K)	m kg s^{-3} K^{-1}
energy density	joule per cubic metre	J/m^3	m^{-1} kg s^{-2}
electric field strength	volt per metre	V/m	m kg s^{-3} A^{-1}
electric charge density	coulomb per cubic metre	C/m^3	m^{-3} s A
electric flux density	coulomb per square metre	C/m^2	m^{-2} s A
permittivity	farad per metre	F/m	m^{-3} kg^{-1} s^4 A^2
permeability	henry per metre	H/m	m kg s^{-2} A^{-2}
molar energy	joule per mole	J/mol	m^2 kg s^{-2} mol^{-1}
molar entropy, molar heat capacity	joule per mole kelvin	J/(mol K)	m^2 kg s^{-2} K^{-1} mol^{-1}
exposure (X and γ rays)	coulomb per kilogram	C/kg	kg^{-1} s A
absorbed dose rate	gray per second	Gy/s	m^2 s^{-3}

Table 58. **Measurement > Units***: SI Derived Units (a.k.a., Metric Derived Units).*

[name] Derivation (derived quantity)	[label] Unit Name	Unit Symbol
Area	Square metre	m^2
volume	Cubic meter	m^3
Speed, velocity	Meter per second	m/s
acceleration	Metre per second squared	m/s^2
Wave number	1 per meter	m^{-1}
Density, mass density	Kilogram per cubic meter	kg/m^3
Specific volume	Cubic meter per kilogram	kg/m^3
Current density	Ampere per square meter	A/m^2
Magnetic field strength	Ampere per meter	A/m
Concentration (of amount of substance)	Mole per cubic meter	mol/m^3
luminance	Candela per square meter	cd/m^2

TABLES

Table 59. Measurement > Units: *Examples of SI derived units formed by using the radian and sterdian.*

Quantity	Unit Name	Unit Symbol
angular velocity	radian per second	rad/s
angular acceleration	radian per second squared	rad/s^2
radiant intensity	watt per steradian	W/sr
radiance	watt per square metre steradian	$W\ m^{-2}\ sr^{-1}$

Table 60. Measurement > Units: *The seven defining constants of the new SI and the corresponding units they define.*

Defining constant	Symbol	Numerical value	Unit
Hyperfine splitting of caesium	Δν (133Cs)hfs	9,192,631,770	$Hz = s^{-1}$
Speed of light in vacuum	c	299,792,458	$Hz = s^{-1}$
Planck constant	h	$6.626070040 \times 10^{-34}$	$J\ s = kg\ m^2\ s^{-1}$
Elementary charge	e	$1.6021766208 \times 10^{-19}$	C = A s
Boltzmann constant	k	$1.38064852 \times 10^{-23}$	$J\ K^{-1} = kg\ m^2$ s−2 K^{-1}
Avogadro constant	N_A	$6.022140857 \times 10^{23}$	mol^{-1}
Luminous efficacy	K_{cd}	683	$cd\ sr\ W^{-1} = cd\ sr\ kg^{-1}\ m^{-2}\ s^3$
The numerical values are taken from the 2014 CODATA adjustment without the present associated uncertainties (not applicable to Δν(133Cs)hfs and c) and may slightly change by 2018.			

Table 61. Measurement > Units: *Physical units as mechanics.*

Physical units: Mechanics			
Quantity and Definition	**Metric cgs**	**Metric MKS**	**English PFS**
Time	Second	Second	Second
Length	Centimeter	Meter	Foot
Mass	Gram	Kilogram	Slug
Velocity v=d/t	centimeter/second	meter/second	foot/ second
Acceleration a = v/t	$centimeter/second^2$	$meter/second^2$	food/ $second^2$
Force F = ma	gm x cm/ sec^2 = dyne	kg x $meter/sec^2$ = newton	Pound
Energy (Work) W = fd	gm x cm2/ sec^2 = erg	kg x $meter^2$/ sec^2 = newton	foot x pound
Power P = W/t	erg/sec	joule/sec = watt	foot x pound/ second
Momentum P = mv	gm x sec = dyne x cm	kg x meter/ sec = N x s	slug x foot/ second
Torque G = Fr	dyne x cm	newton x meter	pound x foot
Frequency	1/sec = hertz	1/sec = hertz	1/sec = hertz

TABLES

Table 62. Measurement > Units*: SI derived units with special names.*

[name] Derivation (derived quantity)	[label] Unit name	Unit Symbol	Expression in terms of other units	Expression in terms of SI base units
plane angle[b]	radian	rad		$m \cdot m^{-1} = 1$
solid angle[b]	Steradian	Sr		$m^2 \cdot m^{-2} = 1$
frequency	Hertz	Hz		s^{-1}
force	newton	N		m kg s^{-2}
pressure, stress	Pascal	Pa	N/m^2	m-1 kg s^{-2}
energy, work quantity of heat	Joule	J	N m	m^2 kg s^{-2}
power, radiant flux	Watt	W	J/s	m2 kg s^{-3}
electric charge, quantity of electricity	Coulomb	C		s A
electric potential, potential difference, electromotive force	volt	V	W/A	m^2 kg s^{-3} A^{-1}
capacitance	farad	F	C/V	m^{-2} kg^{-1} s^4 A^2
electric resistance	ohm	Omega	V/A	m^2 kg s^{-3} A^{-2}
electric conductance	Siemens	S	A/V	m^{-2} kg^{-1} s^3 A^2
magnetic flux	Weber	Wb	V s	m^2 kg s^{-2} A^{-1}
magnetic flux density	Tesla	T	Wb/m2	kg s^{-2} A^{-1}
inductance	Henry	H	Wb/A	m^2 kg s^{-2} A^{-2}
Celsius temperature	Degree Celsius	*C		K
luminous flux	Lumen	Lm	Cd sr	cd $\cdot m^2 \cdot m^{-2}$ = cd
illuminance	Lux	Lx	Lm/m2	cd $\cdot m^2 \cdot m^{-4}$ = cd $\cdot m^{-2}$
activity (of radionuclide)	Becquerel	Bq		s^{-1}
absorbed dose specific energy imparted, kerma	Gray	GY	J/kg	m^2 s^{-2}
dose equivalent	Sievert	Sv	J/kg	m^2 s^{-2}

Table 63. Measurement > Units*: Table of common unit systems.*

[Fundamental] Units in system	[Fundamental] Dimensions of system	Common name of system
Foot-pound-second (FPS)	Length-mass-time	English “system”
Foot-slug-second (FSS)	Length-mass-time	English “system”
Centimeter-gram-second (CGS)	Length-mass-time	Mechanical system
Meter-kilogram-second (MKS)	Length-mass-time	Mechanical system
Meter-Kilogram-second-ampere-kelvin-candela-mole	Length-mass-time-current-temperature-illumiination-amount-of-substance	SI
Meter-Kilogram-second-ampere-kelvin-candela-mole	Length-mass-time-current-temperature-illumiination-amount-of-substance	SI

TABLES

Table 65. Measurement > Units*: SI Units.*

Base Quantity	Base Unit	Symbol [for dimension]	Current SI constants	New SI constants
time	second	s	hyperfine splitting in Cesium-133	same as current SI
length	metre	m	speed of light in vacuum, c	same as current SI
mass	kilogram	kg	mass of international prototype kilogram (IPK)	Planck's constant, h
electric current	Ampere	A	permeability of free space, permittivity of free space	charge of the electron, e
temperature	Kelvin	K	triple point of water, absolute zero	Boltzmann's constant, k
amount of substance	mole	mol	molar mass of Carbon-12	Avogadro constant N_A
luminous intensity	candela	cd	luminous efficacy of a 540 THz source	same as current SI

Table 64. Measurement > Units*: Distance as US and Metric units systems.*

United States System	Metric System
1 mile = 5280 feet	1 kilometer = 1000 meter
1 mile = 1760 yards	1 hectometer = 100 meter
1 rod = 5.5 yards	1 dekameter = 10 meters
1 yard = 3 feet	1 decimeter = 1/10 meter
1 foot = 12 inches	1 centimeter = 1/100 meter

Table 67. Measurement > Unit > Function > Temperature*: Temperatures in celcius and kelvin for important states.*

Name (description)	Celsius	Kelvin
Absolute zero	-273.15 C	0 K
Freezing point of water	0 C	273.15 K
Avg. body temperature	37 C	310.15 K
Boiling point	100 C	373.15 K

Table 66. Measurement > Units*: Derived units.*

Name of quantity	Formula	Derived units
Area	length x breadth	metre-square (m^2)
Volume	length x breadth x height	metre-cubed (m^3)
Speed	distance/time	metre per second ($m\ s^{-1}$)
Pressure	Force/Area	Newton per metre squared (Nm^{-2}) Pascal (Pa)

Table 68. Measurement > Number*: Table showing type of number and its decimal representation.*

Type of number	Decimal Representation
Integer	1.00000000000000000000
Non-repeating fraction	0.25000000000000000000
Repeating fraction	0.12312312312312312312
Irrational number	1.41421356237309504880

TABLES

Table 70. Measurement > Dimensionality: *Order of magnitude in (Dimension: Length; Unit Meter).*

Section	Range (m)		Unit	Examples objects
	≥			
Planck length	-	10^{-35}	ℓ	Quantum
Subatomic	-	10^{-15}	am(10^{-18})	Electron
Atomic and cellular	10^{-15}	10^{-12}	fm	Atomic nucleus, proton, neutron
	10^{-12}	10^{-9}	pm	Wavelength of gamma rays and x-rays, hydrogen atom
	10^{-9}	10^{-6}	nm	DNA helix, virus, wavelength of optical spectrum
Human scale	10^{-6}	10^{-3}	μm	Bacterium, fog water droplet, human hair diameter
	10^{-3}	1	mm	Mosquito, golf ball, domestic cat
	10^{0}	10^{3}	m	Human, automobile, whale, buildings
	10^{3}	10^{6}	km	Mount Everest, length of panama canal, trans-siberian railway, large asteroid
Astronomical	10^{6}	10^{9}	Mm	Moon, Earth, one light-second
	10^{9}	10^{12}	Gm	Sun, one light-minute, earth's orbit
	10^{12}	10^{15}	Tm	Orbits of outer plants, solar system
	10^{15}	10^{18}	Pm	One light-year, distance to Proxima Centauri
	10^{18}	10^{21}	Em	Galactic arm
	10^{21}	10^{24}	Zm	Milky way, distance to Andromeda Galaxy
	10^{24}		Ym	Huge-LQG, Hercules Corona Borealis Great Wall, visible universe

Table 69. Measurement > Number: *Number types.*

Name	Symbol	Meaning
Prime		Prime power factor
Composite		Whole subdivision of a count; for example, 6 is a composite number of 2 x 3
Natural	N	0, 1, 2, 3, 4, ... or 1, 2, 3, 4, ... N0 or N1 are sometimes used
Integer	Z	..., −5, −4, −3, −2, −1, 0, 1, 2, 3, 4, 5, ...
Rational, Ratio[nal], Fraction[able]	Q	>Where a and b are integers and b is not 0 >Perfect squares: √4, √9, √16, √25, √36, √49, √64, √81, √100,..., √256,..., √526, √1024,..., √4096,...)
Irrational, Irratio[nal], Non-fraction[able]	I	Decimal expression is: 1.non-terminal 2.non-repeating (no digit pattern to right of decimal)
Real	R	The limit of a convergent sequence of rational numbers
Complex	C	a + bi where a and b are real numbers and i is a formal square root of −1

TABLES

Table 71. Measurement > Statistics*: Measurement scale types.*

Level of measurement (in scale types)	Characteristics			
	Classification	Order	Equal intervals	True zero point
Nominal	Yes	No	No	No
Ordinal	Yes	Yes	No	No
Interval	Yes	Yes	Yes	No
Ratio	Yes	Yes	Yes	Yes

Table 72. Measurement > Statistics*: Measurement scale types.*

Incremental progress	Measure property	Mathematical operators	Advanced operations	Central tendency
Nominal	Classification, membership	=, !=	Grouping	Mode
Ordinal	Comparison, level	>, <	Sorting	Median
Interval	Difference, affinity	+, -	Yardstick	Mean, deviation
Ratio	Magnitude, amount	*, /	Ratio	Geometric mean, Coefficient of variation

Table 73. Measurement > Statistics*: Classification of scales.*

Classification of scales						
Scale	Operation	Examples	Location	Dispersion	Association	Test
Nominal	Equality	Numbering of objects	Mode			Chi-square
Ordinal	Greater or lesser	Hardness of minerals Street numbers Raw scores	Median	Percentiles	Rank-order correlation	Sign test Run test
Interval	Distance	Temperature: Celsius Position, Time	Arithmetic mean	Standard deviation	Product-moment correlation	t-test F-test
Ratio	Ratio	Numerosity (counts) Length, density Position, time Temperature: Kelvin Loudness: sones Brightness: brils	Geometric mean Harmonic mean	Percent variation		

Table 74. Measurement > Statistics*: Only the ratio scale meets the criteria for all four differentiating properties of a scale of measurement.*

Measurement scales	Indicates difference	Indicates direction of difference	Indicates amount of difference	Absolute zero
Nominal	X			
Ordinal	X	X		
Interval	X	X	X	
Ratio	X	X	X	X

TABLES

Table 75. Measurement > Statistics *Classification of measurement scales based on possible mathematical operations.*

Scale type	Description	Operations	Examples
Nominal	A renaming; can establish equivalence.	=	Colours (red, blue); Team members; Stellar spectral types (O,B,A,F,G,...)
Ordinal	Can establish order	= < >	Moh hardness; Rockwell hardness; Beaufort wind scale; Fahrenheit scales
Interval	Can establish meaningful differences	= < > + -	Date, time of day, year, latitude and longitude, centigrade temperature scale
Metric or ratio	Can establish meaningful ratios	= < > + - /	All SI scales (e.g., length, mass); frequency; thermodynamic temperature
Counting or natural	Counts of objects or events, an integer metric scale	= < > + - /	Apples, tires, birthdays

Table 76. Measurement > Statistics*: Measurement scale types*

Scale type	Level of information	Permissible statistics	Admissible scale transformation	Mathematical structure	Corresponding definition of measurement
Nominal (also denoted as categorical)	Equal/not equal	cell count, mode, contingency correlation, Chi-square	One to one (equality (=))	Standard set structure (unordered)	Assignment of numerals based on rules
Partial order	Order among some but not all categories	Cell count, mode, contingency correlation			
Ordinal	Order among all categories	Median, percentiles	Monotonic increasing (order(<))	Totally ordered set	
Interval	Equal intervals	Mean, standard deviation, correlation, regression, analysis of variance	Positive linear (affine)	Affine line	Measurement as quantification
Ratio	Meaningful zero	All statistics permitted for interval scales plus the following: geometric mean, harmonic mean, coefficient of variation, logarithms	Positive similarities (multiplication)	Field	
Absolute	Numerical count of entities in a given category	Mean, standard deviation, correlation, some forms of regression			

Table 77. Measurement > Statistics*: Scale types.*

Scale type	Characterization	Example (generic	Example (SE)
Nominal	Divides the set of objecvts into categories, with no particular ordering among them	Labeling, classification	Naming of programming language, name of defect type
Ordinal	Divides the set of entities into categories that are ordered	Preference, ranking, difficulty	Ranking of failures (as a measure of failure severity)
Interval	Comparing the differences between values is meaningful	Calendar time, temperature (Farenheit, Celcius)	Beginning and end date of activities (as measures of time distance)
Ratio	There is a meaningful "zero" value, and ratios between values are meaningful.	Length, weight, time intervals, absolute temperature (Kelvin)	Lines of code (as measure of attribute "program length/ size")
Absolute	There are no meaningful transformations of values other than identity	Object count	Count (as measure of attribute "number of lines of code")

TABLES

Table 78. Measurement > Numbers: *Number system scale.*

Number System Sub-name	Real world object	Binary (bi-nary)	Quinary (qui-nary)	Decimal	Sexadecimal a.k.a., hexadecimal (hex)	Base	Names for bases number systems
Base	Stones	two	five	ten	sixteen	2	binary
# of designators (symbols)	Sensation of a stone	2	5	10	16	3	ternary
Digits Increasing count (value), and therein, a base symbolic pattern of increasing orders of magnitude [of that count or value]	No stones	0	0	0	0	4	quaternary
	•	1 (2^0)	1 (5^0)	1 (10^0)	1 (16^0)	5	quinary
	• •	10 (2^1)	2	2	2	6	senary
	• • •	11	3	3	3	7	septenary
	• • • •	100 (2^2)	4	4	4	8	octonary
	• • • • •	101	10 (5^1)	5	5	9	nonary
	• • • • • •	110	11	6	6	10	decimal (denary)
	• • • • • • •	111	12	7	7	11	undenary
	• • • • • • • •	1000 (2^3)	13	8	8	12	duodecimal
	• • • • • • • • •	1001	14	9	9	13	tridecimal
	• • • • • • • • • •	1010	20 (5^2)	10 (10^1)	A	14	quattuordecimal
	• • • • • • • • • • •	1011	21	11	B	15	quindecimal
	• • • • • • • • • • • •	1100	22	12	C	16	sexadecimal
	• • • • • • • • • • • • •	1101	23	13	D	17	septendecimal
	• • • • • • • • • • • • • •	1110	24	14	E	18	octodecimal
	-	1111	30 (5^3)	15	F	19	nonadecimal
	-	10000 (2^4)	31	16	10 (16^1)	20	vigesimal

Table 79. Measurement > Language: *Counting in the English and Chinese languages.*

Written as a decimal (and fraction)	Expression with placement (English). Note, the following words all mean the same thing: "decimal"; "point", and "and".	Expression without placement (Chinese)
1.5 (1 5/10)	one decimal [point] five tenths	one decimal [point] five
3.2 (3 2/10)	three decimal two tenths	three decimal two
1.01 (1 1/100)	one point one hundredth	one decimal zero two
4.975 (4 975/1000)	four and nine hundred seventy-five thousandths	four decimal nine seven eight
5.0016 (5 16/10000)	five and sixteen ten thousandths	five decimal zero zero one six

TABLES

Table 80. <u>**Measurement > Language**</u>*: Linguistic efficiency comparison between numerical written expression in Engels language and Chinese language. The Chinese linguistic expression of numerals is more efficient Some researchers hypothesize that one possible reason some Asian cultures show proficiency in math at an early age ironically has nothing to do with math – it has to do with language. It is easier to learn to count in Chinese than it is in English because it requires learning fewer words.*

Numeral	"English" language		"Chinese" language	
1	one	Ten unique English words	one	Ten unique Chinese words
2	two		two	
3	three		three	
4	four		four	
5	five		five	
6	six		six	
7	seven		seven	
8	eight		eight	
9	nine		nine	
10	ten		ten	
11	eleven	Ten more unique words (total is 20 words)	ten one (or, one ten)	No more unique words (total is 10 words)
12	twelve		ten two	
13	thirteen		ten three	
14	fourteen		ten four	
15	fifteen		ten five	
16	sixteen		ten six	
17	seventeen		ten seven	
18	eighteen		ten eight	
19	nineteen		ten nine	
20	twenty		two ten	
21	twenty one	Eight more unique words (total is 28 words)	two ten one	One more unique word (total is 11 words)
30	thirty		three ten	
40	forty		four ten	
50	fifty		five ten	
60	sixty		six ten	
70	seventy		seven ten	
80	eighty		eight ten	
90	ninety		nine ten	
100	one hundred		one hundred	

TABLES

Table 81. Measurement > Metrology > Semiotics*: Measurement semiotics.*

Percept-ion	Symbols (digits/ letters)	De-notation (numeral/word, numerical signifier)	Con-notation (number/idea)
Mathematics	1 2 3 are digits	153/one hundred fifty three	Visual of a 153 amount
Linguistics	d o g are letters	dog	Visual of a dog
Issue[r]	Identifier	Length	Existent
Mathemat[ics]	123	153/one hundred fifty three	Visual of a 153 amount
Linguist[ics]	dog	dog	Visual of a dog

Table 82. Measurement > Metrology > Properties*: Tabular representation of the measurement of the properties of the objects of model set A. This is 'object oriented' measurement. A class of objects (A) are characterized by the combination of several properties in an object profile (M_1, M_2, m_n).*

Objects of the model set A	Properties			
	M1	M2	•••	mn
a	M_1(a)	M_2(a)		M_n(a)
b	M_1(b)	M_2(b)		M_n(b)
•	•	•		•
•	•	•		•
z	M_1(z)	M_2(z)		M_n(z)

Table 83. Measurement > Method*: Measuring objective and subjective quality-of-life [indicators] based on a focus and method for recording, and then using to predict future, measurement.*

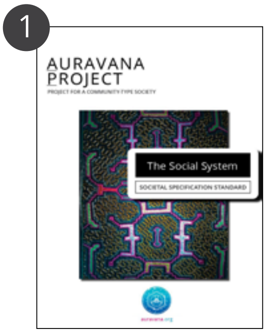
		Intentional Focus of Measurement (is estimation; main criterion)	
		Objective as focused on external non-feelings	*Subjective as focused on feelings*
Method of Measurement (is estimation; subsidiary criterion)	*Objective as external measurement/ estimation*	Focus on external and estimated non-feelings; clearly OWB	-
	Subjective as using subject's self-report	Feelings and other self-reporting data can be objectively studied by externals	Clearly SWB

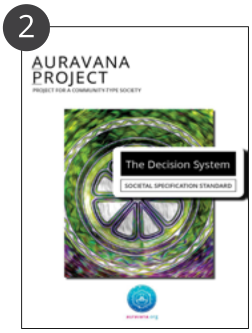
The Auravana Project exists to co-create the emergence of a community-type society through the openly shared development and operation of a information standard, from which is expressed a network of integrated city systems, within which purposefully driven individuals are fulfilled in their development toward a higher potential life experience for themselves and all others. Significant project deliverables include: a societal specification standard and a highly automated, tradeless habitat service operation, which together orient humanity toward fulfillment, wellbeing, and sustainability. The Auravana Project societal standard provides the full specification and explanation for a community-type of society.

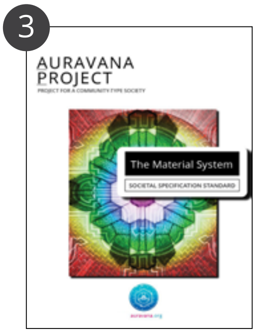
This publication is the Material System for a community-type society. A material system describes the organized structuring of a material environment; the material structuring of community. This material system standard identifies the structures, technologies, and other processes constructed and operated in a material environment, and into a planetary ecology. A material system encodes and expresses our resolved decisions. When a decision resolves into action, that action is specified to occur in the material system. Here, behavior influences the environment, and in turn, the environment influences behavior. The coherent integration and open visualization of the material systems is important if creations are to maintain the highest level of fulfillment for all individuals. This standard represents the encoding of decisions into an environment forming lifestyles within a habitat service system. The visualization and simulation of humanity's connected material integrations is essential for maintaining a set of complex, fulfillment-oriented material constructions. As such, the material system details what has been, what is, and what could be constructed [from our information model] into our environment. This specification depicts, through language and symbols, visualization, and simulation, a material environment consisting of a planetary ecology and embedded network of integrated city systems. For anything that is to be constructed in the material system, there is a written part, a drawing part, and a simulation part, which is also how the material system is sub-divided.

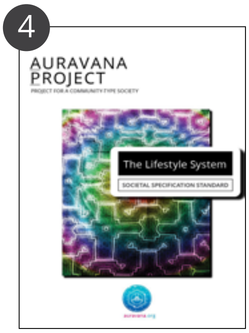
Fundamentally, this standard facilitates individual humans in becoming more aware of who they really are.

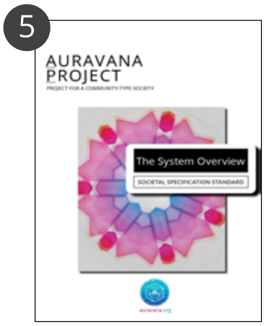
All volumes in the societal standard:

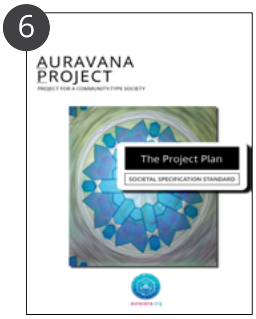












auravana.org

$12.00
ISBN 978-1-7330651-5-3
51200>
9 781733 065153