Abstract: Thermodynamics is a physical branch of science that governs the thermal behavior of dynamical systems from those as simple as refrigerators to those as complex as our expanding universe. The laws of thermodynamics involving conservation of energy and nonconservation of entropy are, without a doubt, two of the most useful and general laws in all sciences. The first law of thermodynamics, according to which energy cannot be created or destroyed, merely transformed from one form to another, and the second law of thermodynamics, according to which the usable energy in an adiabatically isolated dynamical system is always diminishing in spite of the fact that energy is conserved, have had an impact far beyond science and engineering. In this paper, we trace the history of thermodynamics from its classical to its postmodern forms, and present a tutorial and didactic exposition of thermodynamics as it pertains to some of the deepest secrets of the universe.

Keywords: classical thermodynamics; statistical thermodynamics; system thermodynamics; entropy; dynamical systems theory; relativity theory; gravity; life, consciousness, and death; entropic arrow of time

1. Introduction

Even though thermodynamics has provided the foundation for speculation about some of science's most puzzling questions concerning the beginning and the end of the universe, the development of thermodynamics grew out of steam tables and the desire to design and build efficient heat engines, with many scientists and mathematicians expressing concerns about the completeness and clarity of its mathematical foundation over its long and tortuous history. Indeed, many formulations of classical thermodynamics, especially most textbook presentations, poorly amalgamate physics with rigorous mathematics and have had a hard time finding a balance between nineteenth century steam and heat engine engineering, and twenty first century science and mathematics [1–3].

In fact, no other discipline in mathematical science is riddled with so many logical and mathematical inconsistencies, differences in definitions, and ill-defined notation as classical thermodynamics. With a notable few exceptions, more than a century of mathematicians have turned away in disquietude from classical thermodynamics, often overlooking its grandiose unsubstantiated claims and allowing it to slip into an abyss of ambiguity.

The development of the theory of thermodynamics followed two conceptually rather different lines of thought. The first (historically), known as classical thermodynamics, is based on fundamental laws that are assumed as axioms, which in turn are based on experimental evidence. Conclusions are subsequently drawn from them using the notion of a thermodynamic state of a system, which includes temperature, volume, and pressure, among others.

The second, known as statistical thermodynamics, has its foundation in classical mechanics. However, since the state of a dynamical system in mechanics is completely specified point-wise in time by each point-mass position and velocity and since thermodynamic systems contain large numbers of particles (atoms or molecules, typically on the order of $10^{23}$), an ensemble average of different configurations of molecular motion is considered as the state of the system. In this case,
the equivalence between heat and dynamical energy is based on a kinetic theory interpretation reducing all thermal behavior to the statistical motions of atoms and molecules. In addition, the second law of thermodynamics has only statistical certainty wherein entropy is directly related to the relative probability of various states of a collection of molecules.

The second law of thermodynamics is intimately connected to the irreversibility of dynamical processes. In particular, the second law asserts that a dynamical system undergoing a transformation from one state to another cannot be restored to its original state and at the same time restore its environment to its original condition. That is, the status quo cannot be restored everywhere. This gives rise to an increasing quantity known as entropy.

Entropy permeates the whole of nature, and unlike energy, which describes the state of a dynamical system, entropy is a measure of change in the status quo of a dynamical system. Hence, the law that entropy always increases, the second law of thermodynamics, defines the direction of time flow and shows that a dynamical system state will continually change in that direction and thus inevitably approach a limiting state corresponding to a state of maximum entropy. It is precisely this irreversibility of all dynamical processes connoting the running down and eventual demise of the universe that has led writers, historians, philosophers, and theologians to ask profound questions such as: How is it possible for life to come into being in a universe governed by a supreme law that impedes the very existence of life?

Thermodynamics is universal, and hence, in principle, it applies to everything in nature—from simple engineering systems to complex living organisms to our expanding universe. The laws of thermodynamics form the theoretical underpinning of diverse disciplines such as biology, chemistry, climatology, ecology, economics, engineering, genetics, geology, neuroscience, physics, physiology, sociology, and cosmology, and they play a key role in the understanding of these disciplines [4–7]. Modeling the fundamental dynamic phenomena of these disciplines gives rise to large-scale complex [8] dynamical systems that have numerous input, state, and output properties related to conservation, dissipation, and transport of mass, energy, and information. These systems are governed by conservation laws (e.g., mass, energy, fluid, bit, etc.) and are comprised of multiple subsystems or compartments which exchange variable quantities of material via intercompartmental flow laws, and can be characterized as network thermodynamic (i.e., advection-diffusion) systems with compartmental masses, energies, or information playing the role of heat energy in subsystems at different temperatures.

In particular, large-scale compartmental models have been widely used in biology, pharmacology, and physiology to describe the distribution of a substance (e.g., biomass, drug, radioactive tracer, etc.) among different tissues of an organism. In this case, a compartment represents the amount of the substance inside a particular tissue and the intercompartmental flows are due to diffusion processes. In engineering and the physical sciences, compartments typically represent the energy, mass, or information content of the different parts of the system, and different compartments interact by exchanging heat, work energy, and matter.

In ecology and economics, compartments can represent soil and debris, or finished goods and raw materials in different regions, and the flows are due to energy and nutrient exchange (e.g., nitrates, phosphates, carbon, etc.), or money and securities. Compartmental systems can also be used to model chemical reaction systems [7]. In this case, the compartments would represent quantities of different chemical substances contained within the compartment, and the compartmental flows would characterize transformation rates of reactants to products.

The underlying thread of the aforementioned disciplines is the universal principles of conservation of energy and nonconservation of entropy leading to thermodynamic irreversibility, and thus, imperfect animate and inanimate mechanisms—from the diminutive cells in our bodies to the cyclopean spinning galaxies in the heavens. These universal principles underlying the most perplexing secrets of the cosmos are entrusted to the métier of classical thermodynamics—a phenomenological scientific discipline characterized by a purely empirical foundation.
In [3] and more recently in [9], we combined the two universalisms of thermodynamics and dynamical systems theory under a single umbrella, with the latter providing the ideal language for the former, to provide a system-theoretic foundation of thermodynamics and establish rigorous connections between the arrow of time, irreversibility, and the second law of thermodynamics for non-equilibrium systems. Given the proposed dynamical systems formalism of thermodynamics, a question that then arises is whether the proposed dynamical systems framework can be used to shed any new insights into some of the far reaching consequences of the laws of thermodynamics; consequences involving living systems and large-scale physics (i.e., cosmology), the emergence of damping in conservative time-reversible microscopic dynamics, and the ravaging evolution of the universe and, hence, the ultimate destiny of mankind.

Thermodynamicists have always presented theories predicated on the second law of thermodynamics employing equilibrium or near equilibrium models in trying to explain some of the most perplexing secrets of science from a high-level systems perspective. These theories utilize thermodynamic models that include an attempt in explaining the mysteries of the origins of the universe and life, the subsistence of life, biological growth, ecosystem development and sustenance, and biological evolution, as well as explaining nonliving organized systems such as galaxies, stars, accretion disks, black holes, convection cells, tornados, hurricanes, eddies, vortices, and river deltas, to name but a few examples. However, most of these thermodynamic models are restricted to equilibrium or near equilibrium systems, and hence, are inappropriate in addressing some of these deep scientific mysteries as they involve nonequilibrium nonlinear dynamical processes undergoing temporal and spatial changes.

Given that dynamical systems theory has proven to be one of the most universal formalisms in describing manifestations of nature which involve change and time, it provides the ideal framework for developing a postmodern foundation for nonequilibrium thermodynamics [3,9]. This can potentially lead to mathematical models with greater precision and self-consistency; and while “self-consistency is not necessarily truth, self-inconsistency is certainly falsehood” [10].

The proposed dynamical systems framework of thermodynamics given in [9] can potentially provide deeper insights into some of most perplexing questions concerning the origins and fabric of our universe that require dynamical system models that are far from equilibrium. In addition, dynamical thermodynamics can foster the development of new frameworks in explaining the fundamental thermodynamic processes of nature, explore new hypotheses that challenge the use of classical thermodynamics, and allows for the development of new assertions that can provide deeper insights into the constitutive mechanisms that describe acute microcosms and macrocosms.

In this paper, we trace the long and tortuous history of thermodynamics and present a plenary exposition on how system thermodynamics, as developed in [9], can address some of science’s most puzzling questions concerning dynamical processes occurring in nature. Among these processes are the thermodynamics of living systems, the origins of life and the universe, consciousness, health, and death. The nuances underlying these climacterical areas are the subject matter of this paper. We stress, however, that unlike the development in [9], which provides a rigorous mathematical presentation of the key concepts of system thermodynamics, in this paper we give a high-level scientific discussion of these very important subjects.

2. An Overview of Classical Thermodynamics

Energy is a concept that underlies our understanding of all physical phenomena and is a measure of the ability of a dynamical system to produce changes (motion) in its own system state as well as changes in the system states of its surroundings. Thermodynamics is a physical branch of science that deals with laws governing energy flow from one body to another and energy transformations from one form to another. These energy flow laws are captured by the fundamental principles known as the first and second laws of thermodynamics. The first law of thermodynamics gives a precise formulation of the equivalence between heat and work and states that among all system transformations, the net
system energy is conserved. Hence, energy cannot be created out of nothing and cannot be destroyed; it can merely be transferred from one form to another.

The law of conservation of energy is not a mathematical truth, but rather the consequence of an immeasurable culmination of observations over the chronicle of our civilization and is a fundamental axiom of the science of heat. The first law does not tell us whether any particular process can actually occur, that is, it does not restrict the ability to convert work into heat or heat into work, except that energy must be conserved in the process. The second law of thermodynamics asserts that while the system energy is always conserved, it will be degraded to a point where it cannot produce any useful work. More specifically, for any cyclic process that is shielded from heat exchange with its environment, it is impossible to extract work from heat without at the same time discarding some heat, giving rise to an increasing quantity known as entropy.

While energy describes the state of a dynamical system, entropy refers to changes in the status quo of the system and is associated with disorder and the amount of wasted energy in a dynamical (energy) transformation from one state (form) to another. Since the system entropy increases, the entropy of a dynamical system tends to a maximum, and thus time, as determined by system entropy increase [11–13], flows on in one direction only. Even though entropy is a physical property of matter that is not directly observable, it permeates the whole of nature, regulating the arrow of time, and is responsible for the enfeeblement and eventual demise of the universe [14,15]. While the laws of thermodynamics form the foundation to basic engineering systems, chemical reaction systems, nuclear explosions, cosmology, and our expanding universe, many mathematicians and scientists have expressed concerns about the completeness and clarity of the different expositions of thermodynamics over its long and tortuous history; see [2,16–23].

Since the specific motion of every molecule of a thermodynamic system is impossible to predict, a macroscopic model of the system is typically used, with appropriate macroscopic states that include pressure, volume, temperature, internal energy, and entropy, among others. One of the key criticisms of the macroscopic viewpoint of thermodynamics, known as classical thermodynamics, is the inability of the model to provide enough detail of how the system really evolves; that is, it is lacking a kinetic mechanism for describing the behavior of heat.

In developing a kinetic model for heat and dynamical energy, a thermodynamically consistent energy flow model should ensure that the system energy can be modeled by a diffusion equation in the form of a parabolic partial differential equation or a divergence structure first-order hyperbolic partial differential equation arising in models of conservation laws. Such systems are infinite-dimensional, and hence, finite-dimensional approximations are of very high order, giving rise to large-scale dynamical systems with macroscopic energy transfer dynamics. Since energy is a fundamental concept in the analysis of large-scale dynamical systems, and heat (energy) is a fundamental concept of thermodynamics involving the capacity of hot bodies (more energetic subsystems) to produce work, thermodynamics is a theory of large-scale dynamical systems.

High-dimensional dynamical systems can arise from both macroscopic and microscopic points of view. Microscopic thermodynamic models can have the form of a distributed-parameter model or a large-scale system model comprised of a large number of interconnected Hamiltonian subsystems. For example, in a crystalline solid every molecule in a lattice can be viewed as an undamped vibrational mode comprising a distributed-parameter model in the form of a second-order hyperbolic partial differential equation. In contrast to macroscopic models involving the evolution of global quantities (e.g., energy, temperature, entropy, etc.), microscopic models are based upon the modeling of local quantities that describe the atoms and molecules that make up the system and their speeds, energies, masses, angular momenta, behavior during collisions, etc. The mathematical formulations based on these quantities form the basis of statistical mechanics.

Thermodynamics based on statistical mechanics is known as statistical thermodynamics and involves the mechanics of an ensemble of many particles (atoms or molecules) wherein the detailed description of the system state loses importance and only average properties of large numbers
of particles are considered. Since microscopic details are obscured on the macroscopic level, it is appropriate to view a macroscopic model as an inherent model of uncertainty. However, for a thermodynamic system the macroscopic and microscopic quantities are related since they are simply different ways of describing the same phenomena. Thus, if the global macroscopic quantities can be expressed in terms of the local microscopic quantities, then the laws of thermodynamics could be described in the language of statistical mechanics.

This interweaving of the microscopic and macroscopic points of view leads to diffusion being a natural consequence of dimensionality and, hence, uncertainty on the microscopic level, despite the fact that there is no uncertainty about the diffusion process per se. Thus, even though as a limiting case a second-order hyperbolic partial differential equation purports to model an infinite number of modes, in reality much of the modal information (e.g., position, velocity, energies, etc.) are only poorly known, and hence, such models are largely idealizations. With increased dimensionality comes an increase in uncertainty leading to a greater reliance on macroscopic quantities so that the system model becomes more diffusive in character.

Thermodynamics was spawned from the desire to design and build efficient heat engines, and it quickly spread to speculations about the universe upon the discovery of entropy as a fundamental physical property of matter. The theory of classical thermodynamics was predominantly developed by Carnot, Clausius, Kelvin, Planck, Gibbs, and Carathéodory [24], and its laws have become one of the most firmly established scientific achievements ever accomplished. The pioneering work of Carnot [25] was the first to establish the impossibility of a perpetuum mobile of the second kind [26] by constructing a cyclical process (now known as the Carnot cycle) involving four thermodynamically reversible processes operating between two heat reservoirs at different temperatures, and showing that it is impossible to extract work from heat without at the same time discarding some heat.

Carnot’s main assumption (now known as Carnot’s principle) was that it is impossible to perform an arbitrarily often repeatable cycle whose only effect is to produce an unlimited amount of positive work. In particular, Carnot showed that the efficiency of a reversible [27–29] cycle—that is, the ratio of the total work produced during the cycle and the amount of heat transferred from a boiler (furnace) to a cooler (refrigerator)—is bounded by a universal maximum, and this maximum is only a function of the temperatures of the boiler and the cooler, and not on the nature of the working substance.

Both heat reservoirs (i.e., furnace and refrigerator) are assumed to have an infinite source of heat so that their state is unchanged by their heat exchange with the engine (i.e., the device that performs the cycle), and hence, the engine is capable of repeating the cycle arbitrarily often. Carnot’s result (now known as Carnot’s theorem) was remarkably arrived at using the erroneous concept that heat is an indestructible substance, that is, the caloric theory of heat [30]. This theory of heat was proposed by Black and was based on the incorrect assertion that the temperature of a body was determined by the amount of caloric that it contained; an imponderable, indestructible, and highly elastic fluid that surrounded all matter and whose self repulsive nature was responsible for thermal expansion.

Different notions of the conservation of energy can be traced as far back to the ancient Greek philosophers Thales (~624–~546 B.C.), Herakleitos (~535–~475 B.C.), and Empedocles (~490–~430 B.C.). Herakleitos postulates that nothing in nature can be created out of nothing, and nothing that disappears ceases to exist [31], whereas Empedocles asserts that nothing comes to be or perishes in nature [32]. The mechanical equivalence principle of heat and work energy in its modern form, however, was developed by many scientists in the nineteenth century. Notable contributions include the work of Mayer, Joule, Thomson (Lord Kelvin), Thompson (Count Rumford), Helmholtz, Clausius, Maxwell, and Planck.

Even though many scientists are credited with the law of conservation of energy, it was first discovered independently by Mayer and Joule. Mayer—a surgeon—was the first to state the mechanical equivalence of heat and work energy in its modern form after noticing that his patients’ blood in the tropics was a deeper red leading him to deduce that they were consuming less oxygen, and hence
less energy, in order to maintain their body temperature in a hotter climate. This observation in slower human metabolism along with the link between the body’s heat release and the chemical energy released by the combustion of oxygen led Mayer to the discovery that heat and mechanical work are interchangeable.

Joule was the first to provide a series of decisive, quantitative studies in the 1840’s showing the equivalence between heat and mechanical work. Specifically, he showed that if a thermally isolated system is driven from an initial state to a final state, then the work done is only a function of the initial and final equilibrium states, and not dependent on the intermediate states or the mechanism doing the work. This path independence property along with the irrelevancy by what method the work was done lead to the definition of the internal energy function as a new thermodynamic coordinate characterizing the state of a thermodynamic system. In other words, heat or work do not contribute separately to the internal energy function; only the sum of the two matters.

Using a macroscopic approach and building on the work of Carnot, Clausius [33–36] was the first to introduce the notion of entropy as a physical property of matter and establish the two main laws of thermodynamics involving conservation of energy and nonconservation of entropy [37]. Specifically, using conservation of energy principles, Clausius showed that Carnot’s principle is valid. Furthermore, Clausius postulated that it is impossible to perform a cyclic system transformation whose only effect is to transfer heat from a body at a given temperature to a body at a higher temperature. From this postulate Clausius established the second law of thermodynamics as a statement about entropy increase for adiabatically isolated systems (i.e., systems with no heat exchange with the environment).

From this statement Clausius goes on to state what have become known as the most controversial words in the history of thermodynamics and perhaps all of science; namely, the entropy of the universe is tending to a maximum, and the total state of the universe will inevitably approach a limiting state. Clausius’ second law decrees that the usable energy in the universe is locked towards a path of degeneration, sliding towards a state of quietus. The fact that the entropy of the universe is a thermodynamically undefined concept led to serious criticism of Clausius’ grand universal generalizations by many of his contemporaries as well as numerous scientists, natural philosophers, and theologians who followed.

Clausius’ concept of the universe approaching a limiting state was inadvertently based on an analogy between a universe and a finite adiabatically isolated system possessing a finite energy content. His eschatological conclusions are far from obvious for complex dynamical systems with dynamical states far from equilibrium and involving processes beyond a simple exchange of heat and mechanical work. It is not clear where the heat absorbed by the system, if that system is the universe, needed to define the change in entropy between two system states comes from. Nor is it clear whether an infinite and endlessly expanding universe governed by the theory of general relativity has a final equilibrium state.

An additional caveat is the delineation of energy conservation when changes in the curvature of spacetime need to be accounted for. In this case, the energy density tensor in Einstein’s field equations is only covariantly conserved (i.e., locally conserved in free falling coordinates) since it does not account for gravitational energy—an unsolved problem in the general theory of relativity. In particular, conservation of energy and momentum laws, wherein a global time coordinate does not exist, has lead to one of the fundamental problems in general relativity. Specifically, in general relativity involving a curved spacetime (i.e., a semi-Riemannian spacetime), the action of the gravitational field is invariant with respect to arbitrary coordinate transformations in semi-Riemannian spacetime with a nonvanishing Jacobian containing a large number of Lie groups.

In this case, it follows from Noether’s theorem [38,39], which derives conserved quantities from symmetries and states that every differentiable symmetry of a dynamical action has a corresponding conservation law, that a large number of conservation laws exist; some of which are not physical. In contrast, the classical conservation laws of physics, which follow from time translation invariance, are determined by an invariant property under a particular Lie group with the
conserved quantities corresponding to the parameters of the group. In special relativity, conservation of energy and momentum is a consequence of invariance through the action of infinitesimal translation of the inertial coordinates, wherein the Lorentz transformation relates inertial systems in different inertial coordinates.

In general relativity the momentum-energy equivalence principle holds only in a local region of spacetime—a flat or Minkowski spacetime. In other words, the energy-momentum conservation laws in gravitation theory involve gauge conservation laws with local time transformations, wherein the covariant transformation generators are canonical horizontal prolongations of vector fields on a world manifold [40], and hence, in a curved spacetime there does not exist a global energy-momentum conservation law. Nevertheless, the law of conservation of energy is as close to an absolute truth as our incomprehensible universe will allow us to deduce. In his later work [36], Clausius remitted his famous claim that the entropy of the universe is tending to a maximum.

In parallel research Kelvin [41,42] developed similar, and in some cases identical, results as Clausius, with the main difference being the absence of the concept of entropy. Kelvin’s main view of thermodynamics was that of a universal irreversibility of physical phenomena occurring in nature. Kelvin further postulated that it is impossible to perform a cyclic system transformation whose only effect is to transform into work heat from a source that is at the same temperature throughout [43–48]. Without any supporting mathematical arguments, Kelvin goes on to state that the universe is heading towards a state of eternal rest wherein all life on Earth in the distant future shall perish. This claim by Kelvin involving a universal tendency towards dissipation has come to be known as the heat death of the universe.

The universal tendency towards dissipation and the heat death of the universe were expressed long before Kelvin by the ancient Greek philosophers Herakleitos and Leukippos (∼480–∼420 B.C.). In particular, Herakleitos states that this universe which is the same everywhere, and which no one god or man has made, existed, exists, and will continue to exist as an eternal source of energy set on fire by its own natural laws, and will dissipate under its own laws [49]. Herakleitos’ profound statement created the foundation for all metaphysics and physics and marks the beginning of science postulating the big bang theory as the origin of the universe as well as the heat death of the universe. A century after Herakleitos, Leukippos declared that from its genesis, the cosmos has spawned multitudinous worlds that evolve in accordance to a supreme law that is responsible for their expansion, enfeeblement, and eventual demise [50].

Building on the work of Clausius and Kelvin, Planck [51,52] refined the formulation of classical thermodynamics. From 1897 to 1964, Planck’s treatise [51] underwent eleven editions and is considered the definitive exposition on classical thermodynamics. Nevertheless, these editions have several inconsistencies regarding key notions and definitions of reversible and irreversible processes [53–55]. Planck’s main theme of thermodynamics is that entropy increase is a necessary and sufficient condition for irreversibility. Without any proof (mathematical or otherwise), he goes on to conclude that every dynamical system in nature evolves in such a way that the total entropy of all of its parts increases. In the case of reversible processes, he concludes that the total entropy remains constant.

Unlike Clausius’ entropy increase conclusion, Planck’s increase entropy principle is not restricted to adiabatically isolated dynamical systems. Rather, it applies to all system transformations wherein the initial states of any exogenous system, belonging to the environment and coupled to the transformed dynamical system, return to their initial condition. It is important to note that Planck’s entire formulation is restricted to homogeneous systems for which the thermodynamical state is characterized by two thermodynamic state variables, that is, a fluid. His formulation of entropy and the second law is not defined for more complex systems that are not in equilibrium and in an environment that is more complex than one comprising a system of ideal gases.

Unlike the work of Clausius, Kelvin, and Planck involving cyclical system transformations, the work of Gibbs [56] involves system equilibrium states. Specifically, Gibbs assumes a thermodynamic state of a system involving pressure, volume, temperature, energy, and entropy, among others, and proposes that
an isolated system [57] (i.e., a system with no energy exchange with the environment) is in equilibrium if and only if all possible variations of the state of the system that do not alter its energy, the variation of the system entropy is negative semidefinite.

Gibbs also proposed a complimentary formulation of his principle involving a principle of minimal energy. Namely, for an equilibrium of any isolated system, it is necessary and sufficient that in all possible variations of the state of the system that do not alter its entropy, the variation of its energy shall either vanish or be positive. Hence, the system energy is minimized at the system equilibrium.

Gibbs’ principles give necessary and sufficient conditions for a thermodynamically stable equilibrium and should be viewed as variational principles defining admissible (i.e., stable) equilibrium states. Thus, they do not provide any information about the dynamical state of the system as a function of time nor any conclusions regarding entropy increase or energy decrease in a dynamical system transformation.

Carathéodory [58,59] was the first to give a rigorous axiomatic mathematical framework for thermodynamics. In particular, using an equilibrium thermodynamic theory, Carathéodory assumes a state space endowed with a Euclidean topology and defines the equilibrium state of the system using thermal and deformation coordinates. Next, he defines an adiabatic accessibility relation wherein a reachability condition of an adiabatic process [58–60] is used such that an empirical statement of the second law characterizes a mathematical structure for an abstract state space. Even though the topology in Carathéodory’s thermodynamic framework is induced on $\mathbb{R}^n$ (the space of $n$-tuples of reals) by taking the metric to be the Euclidean distance function and constructing the corresponding neighborhoods, the metrical properties of the state space do not play a role in his theory as there is no preference for a particular set of system coordinates.

Carathéodory’s postulate for the second law states that in every open neighborhood of any equilibrium state of a system, there exist equilibrium states such that for some second open neighborhood contained in the first neighborhood, all the equilibrium states in the second neighborhood cannot be reached by adiabatic processes from equilibrium states in the first neighborhood. From this postulate Carathéodory goes on to show that for a special class of systems, which he called simple systems, there exists a locally defined entropy and an absolute temperature on the state space for every simple system equilibrium state. In other words, Carathéodory’s postulate establishes the existence of an integrating factor for the heat transfer in an infinitesimal reversible process for a thermodynamic system of an arbitrary number of degrees of freedom that makes entropy an exact (i.e., total) differential.

Unlike the work of Clausius, Kelvin, Planck, and Gibbs, Carathéodory provides a topological formalism for the theory of thermodynamics, which elevates the subject to the level of other theories of modern physics. Specifically, the empirical statement of the second law is replaced by an abstract state space formalism, wherein the second law is converted into a local topological property endowed with a Euclidean metric. This parallels the development of relativity theory, wherein Einstein’s original special theory started from empirical principles—e.g., the velocity of light is invariant in all inertial frames—and then replaced by an abstract geometrical structure; the Minkowski spacetime, wherein the empirical principles are converted into local topological properties of the Minkowski metric. However, one of the key limitations of Carathéodory’s work is that his principle is too weak in establishing the existence of a global entropy function.

Adopting a microscopic viewpoint, Boltzmann [61] was the first to give a probabilistic interpretation of entropy involving different configurations of molecular motion of the microscopic dynamics. Specifically, Boltzmann reinterpreted thermodynamics in terms of molecules and atoms by relating the mechanical behavior of individual atoms with their thermodynamic behavior by suitably averaging properties of the individual atoms. In particular, even though individually each molecule and atom obeys Newtonian mechanics, he used the science of statistical mechanics to bridge between the microscopic details and the macroscopic behavior to try and find a mechanical underpinning of the second law.
Even though Boltzmann was the first to give a probabilistic interpretation of entropy as a measure of the disorder of a physical system involving the evolution towards the largest number of possible configurations of the system’s states relative to its ordered initial state, Maxwell was the first to use statistical methods to understand the behavior of the kinetic theory of gases. In particular, he postulated that it is not necessary to track the positions and velocities of each individual atom and molecule, but rather it suffices to know their position and velocity distributions; concluding that the second law is merely statistical. His distribution law for the kinetic theory of gases describes an exponential function giving the statistical distribution of the velocities and energies of the gas molecules at thermal equilibrium and provides an agreement with classical (i.e., nonquantum) mechanics.

Although the Maxwell speed distribution law agrees remarkably well with observations for an assembly of weakly-interacting particles that are distinguishable, it fails for indistinguishable (i.e., identical) particles at high densities. In these regions, speed distributions predicated on the principles of quantum physics must be used; namely, the Fermi-Dirac and Bose-Einstein distributions. In this case, the Maxwell statistics closely agree with the Bose-Einstein statistics for bosons (photons, α-particles, and all nuclei with an even mass number) and the Fermi-Dirac statistics for fermions (electrons, photons, and neutrons).

Boltzmann, however, further showed that even though individual atoms are assumed to obey the laws of Newtonian mechanics, by suitably averaging over the velocity distributions of these atoms the microscopic (mechanical) behavior of atoms and molecules produced effects visible on a macroscopic (thermodynamic) scale. He goes on to argue that Clausius’ thermodynamic entropy (a macroscopic quantity) is proportional to the logarithm of the probability that a system will exist in the state it is in relative to all possible states it could be in. Thus, the entropy of a thermodynamic system state (macrostate) corresponds to the degree of uncertainty about the actual system mechanical state (microstate) when only the thermodynamic system state (macrostate) is known. Hence, the essence of Boltzmann thermodynamics is that thermodynamic systems with a constant energy level will evolve from a less probable state to a more probable state with the equilibrium system state corresponding to a state of maximum entropy (i.e., highest probability).

Interestingly, Boltzmann’s original thinking on the subject of entropy increase involved nondecreasing of entropy as an absolute certainty and not just as a statistical certainty. In the 1870s and 1880s his thoughts on this matter underwent significant refinements and shifted to a probabilistic viewpoint after interactions with Maxwell, Kelvin, Loschmidt, Gibbs, Poincaré, Burbury, and Zermelo; all of who criticized his original formulation.

In statistical thermodynamics the Boltzmann entropy formula relates the entropy $S$ of an ideal gas to the number of distinct microstates $W$ corresponding to a given macrostate as $S = k \log_e W$, where $k$ is the Boltzmann constant [62]. Even though Boltzmann was the first to link the thermodynamic entropy of a macrostate for some probability distribution of all possible microstates generated by different positions and momenta of various gas molecules [63], it was Planck who first stated (without proof) this entropy formula in his work on black body radiation [64]. In addition, Planck was also the first to introduce the precise value of the Boltzmann constant to the formula; Boltzmann merely introduced the proportional logarithmic connection between the entropy $S$ of an observed macroscopic state, or degree of disorder of a system, to the thermodynamic probability of its occurrence $W$, never introducing the constant $k$ to the formula.

To further complicate matters, in his original paper [64] Planck stated the formula without derivation or clear justification; a fact that deeply troubled Albert Einstein [65]. Despite the fact that numerous physicists consider $S = k \log_e W$ as the second most important formula of physics—second to Einstein’s $E = mc^2$—for its unquestionable success in computing the thermodynamic entropy of isolated systems [66], its theoretical justification remains ambiguous and vague in most statistical thermodynamics textbooks. In this regard, Khinchin [55] (p. 142) writes: “All existing attempts to give a general proof of [Boltzmann’s entropy formula] must be considered as an aggregate of logical and mathematical errors superimposed on a general confusion in the definition of basic quantities”.

$$S = k \log_e W$$
In the first half of the twentieth century, the macroscopic (classical) and microscopic (statistical) interpretations of thermodynamics underwent a long and fierce debate. To exacerbate matters, since classical thermodynamics was formulated as a physical theory and not a mathematical theory, many scientists and mathematical physicists expressed concerns about the completeness and clarity of the mathematical foundation of thermodynamics \([2,19,67]\). In fact, many fundamental conclusions arrived at by classical thermodynamics can be viewed as paradoxical.

For example, in classical thermodynamics the notion of entropy (and temperature) is only defined for equilibrium states. However, the theory concludes that nonequilibrium states transition towards equilibrium states as a consequence of the law of entropy increase! Furthermore, classical thermodynamics is mainly restricted to systems in equilibrium. The second law infers that for any transformation occurring in an isolated system, the entropy of the final state can never be less than the entropy of the initial state. In this context, the initial and final states of the system are equilibrium states. However, by definition, an equilibrium state is a system state that has the property that whenever the state of the system starts at the equilibrium state it will remain at the equilibrium state for all future time unless an exogenous input acts on the system. Hence, the entropy of the system can only increase if the system is not isolated!

Many aspects of classical thermodynamics are riddled with such inconsistencies, and hence it is not surprising that many formulations of thermodynamics, especially most textbook expositions, poorly amalgamate physics with rigorous mathematics. Perhaps this is best eulogized in [2] (p. 6), wherein Truesdell describes the present state of the theory of thermodynamics as a “dismal swamp of obscurity”. In a desperate attempt to try and make sense of the writings of de Groot, Mazur, Casimir, and Prigogine he goes on to state that there is “something rotten in the [thermodynamic] state of the Low Countries” [2] (p. 134).

Brush [19] (p. 581) remarks that “anyone who has taken a course in thermodynamics is well aware, the mathematics used in proving Clausius’ theorem . . . [has] only the most tenuous relation to that known to mathematicians”. Born [68] (p. 119) admits that “I tried hard to understand the classical foundations of the two theorems, as given by Clausius and Kelvin; . . . but I could not find the logical and mathematical root of these marvelous results”. More recently, Arnold [67] (p. 163) writes that “every mathematician knows it is impossible to understand an elementary course in thermodynamics”.

As we have outlined, it is clear that there have been many different presentations of classical thermodynamics with varying hypotheses and conclusions. To exacerbate matters, there are also many vaguely defined terms and functions that are central to thermodynamics such as entropy, enthalpy, free energy, quasistatic, nearly in equilibrium, extensive variables, intensive variables, reversible, irreversible, etc. These functions’ domain and codomain are often unspecified and their local and global existence, uniqueness, and regularity properties are unproven.

Moreover, there are no general dynamic equations of motion, no ordinary or partial differential equations, and no general theorems providing mathematical structure and characterizing classes of solutions. Rather, we are asked to believe that a certain differential can be greater than something which is not a differential defying the syllogism of differential calculus, line integrals approximating adiabatic and isothermal paths result in alternative solutions annulling the fundamental theorem of integral calculus, and we are expected to settle for descriptive and unmathematical wordplay in explaining key principles that have far reaching consequences in engineering, science, and cosmology.

Furthermore, the careless and considerable differences in the definitions of two of the key notions of thermodynamics—namely, the notions of reversibility and irreversibility—have contributed to the widespread confusion and lack of clarity of the exposition of classical thermodynamics over the past one and a half centuries. For example, the concept of reversible processes as defined by Clausius, Kelvin, Planck, and Carathédory have very different meanings. In particular, Clausius defines a reversible (umkehrbar) process as a slowly varying process wherein successive states of this process differ by infinitesimals from the equilibrium system states. Such system transformations are commonly referred to as quasistatic transformations in the thermodynamic literature. Alternatively, Kelvin’s
notions of reversibility involve the ability of a system to completely recover its initial state from the final system state.

Planck introduced several notions of reversibility. His main notion of reversibility is one of complete reversibility and involves recoverability of the original state of the dynamical system while at the same time restoring the environment to its original condition. Unlike Clausius’ notion of reversibility, Kelvin’s and Planck’s notions of reversibility do not require the system to exactly retrace its original trajectory in reverse order. Carathéodory’s notion of reversibility involves recoverability of the system state in an adiabatic process resulting in yet another definition of thermodynamic reversibility. These subtle distinctions of (ir)reversibility are often unrecognized in the thermodynamic literature. Notable exceptions to this fact include [69,70], with [70] providing an excellent exposition of the relation between irreversibility, the second law of thermodynamics, and the arrow of time.

3. Thermodynamics and the Arrow of Time

The arrow of time [71,72] and the second law of thermodynamics is one of the most famous and controversial problems in physics. The controversy between ontological time (i.e., a timeless universe) and the arrow of time (i.e., a constantly changing universe) can be traced back to the famous dialogues between the ancient Greek philosophers Parmenides [73,74] and Herakleitos on being and becoming. Parmenides, like Einstein, insisted that time is an illusion, that there is nothing new, and that everything is (being) and will forever be. This statement is, of course, paradoxical since the status quo changed after Parmenides wrote his famous poem On Nature.

Parmenides maintained that we all exist within spacetime, and time is a one-dimensional continuum in which all events, regardless of when they happen from any given perspective, simply are. All events exist and occupy ordered points in spacetime, and hence, reality envelops past, present, and future equally. More specifically, our picture of the universe at a given moment is identical and contains exactly the same events; we simply have different conceptions of what exists at that moment, and hence, different conceptions of reality. Conversely, the Heraclitan flux doctrine maintains that nothing ever is, and everything is becoming. In this regard, time gives a different ontological status of past, present, and future resulting in an ontological transition, creation, and actualization of events. More specifically, the unfolding of events in the flow of time have counterparts in reality.

Herakleitos’ aphorism is predicated on change (becoming); namely, the universe is in a constant state of flux and nothing is stationary—Ta πάντα ρεί καὶ οὐδὲν μένει. Furthermore, Herakleitos goes on to state that the universe evolves in accordance with its own laws which are the only unchangeable things in the universe (e.g., universal conservation and nonconservation laws). His statements that everything is in a state of flux—Ta πάντα ρεί—and that man cannot step into the same river twice, because neither the man nor the river is the same—Ποταμεία τοίς αὐτοῖς εμβάφομεν τὸ καὶ οὐκ εμβάφομεν, εἶμεν τὲ καὶ οὐκ εἶμεν—give the earliest perception of irreversibility of nature and the universe along with time’s arrow. The idea that the universe is in constant change and there is an underlying order to this change—the Logos (Λόγος)—postulates the existence of entropy as a physical property of matter permeating the whole of nature and the universe.

Herakleitos’ statements are completely consistent with the laws of thermodynamics which are intimately connected to the irreversibility of dynamical processes in nature. In addition, his aphorisms go beyond the worldview of classical thermodynamics and have deep relativistic ramifications to the spacetime fabric of the cosmos. Specifically, Herakleitos’ profound statement—All matter is exchanged for energy, and energy for all matter (Πυρός τε ἀνταμαβή τὰ πάντα καὶ πῦρ ἁπάντων)—is a statement of the law of conservation of mass-energy and is a precursor to the principle of relativity. In describing the nature of the universe Herakleitos postulates that nothing can be created out of nothing, and nothing that disappears ceases to exist. This totality of forms, or mass-energy equivalence, is eternal [75] and unchangeable in a constantly changing universe.
The arrow of time [76] remains one of physics’ most perplexing enigmas [52,77–83]. Even though time is one of the most familiar concepts humankind has ever encountered, it is the least understood. Puzzling questions of time’s mysteries have remained unanswered throughout the centuries [84]. Questions such as, Where does time come from? What would our universe look like without time? Can there be more than one dimension to time? Is time truly a fundamental appurtenance woven into the fabric of the universe, or is it just a useful edifice for organizing our perception of events? Why is the concept of time hardly ever found in the most fundamental physical laws of nature and the universe? Can we go back in time? And if so, can we change past events?

Human experience perceives time flow as unidirectional; the present is forever flowing towards the future and away from a forever fixed past. Many scientists have attributed this emergence of the direction of time flow to the second law of thermodynamics due to its intimate connection to the irreversibility of dynamical processes [85]. In this regard, thermodynamics is disjoint from Newtonian and Hamiltonian mechanics (including Einstein’s relativistic and Schrödinger’s quantum extensions), since these theories are invariant under time reversal, that is, they make no distinction between one direction of time and the other. Such theories possess a time-reversal symmetry, wherein, from any given moment of time, the governing laws treat past and future in exactly the same way [86–89]. It is important to stress here that time-reversal symmetry applies to dynamical processes whose reversal is allowed by the physical laws of nature, not a reversal of time itself. It is irrelevant of whether or not the reversed dynamical process actually occurs in nature; it suffices that the theory allows for the reversed process to occur.

The simplest notion of time-reversal symmetry is the statement wherein the physical theory in question is time-reversal symmetric in the sense that given any solution \( x(t) \) to a set of dynamic equations describing the physical laws, then \( x(-t) \) is also a solution to the dynamic equations. For example, in Newtonian mechanics this implies that there exists a transformation \( R(q,p) \) such that \( R(q,p) \circ x(t) = x(-t) \circ R(q,p) \), where \( \circ \) denotes the composition operator and \( x(-t) = [q(-t), -p(-t)]^T \) represents the particles that pass through the same position as \( q(t) \), but in reverse order and with reverse velocity \(-p(-t)\). It is important to note that if the physical laws describe the dynamics of particles in the presence of a field (e.g., an electromagnetic field), then the reversal of the particle velocities is insufficient for the equations to yield time-reversal symmetry. In this case, it is also necessary to reverse the field, which can be accomplished by modifying the transformation \( R \) accordingly.

As an example of time-reversal symmetry, a film run backwards of a harmonic oscillator over a full period or a planet orbiting the Sun would represent possible events. In contrast, a film run backwards of water in a glass coalescing into a solid ice cube or ashes self-assembling into a log of wood would immediately be identified as an impossible event. Over the centuries, many philosophers and scientists shared the views of a Parmenidean frozen river time theory. However, since the advent of the science of thermodynamics in the nineteenth century, philosophy and science took a different point of view with the writings of Hegel, Bergson, Heidegger, Clausius, Kelvin, and Boltzmann; one involving time as our existential dimension. The idea that the second law of thermodynamics provides a physical foundation for the arrow of time has been postulated by many authors [3,71,77,90–94]. However, a convincing mathematical argument of this claim has never been given [70,78,82].

The complexities inherent with the afore statement are subtle and are intimately coupled with the universality of thermodynamics, entropy, gravity, and cosmology (see Section 9). A common misconception of the principle of the entropy increase is surmising that if entropy increases in forward time, then it necessarily decreases in backward time. However, entropy and second law do not alter the laws of physics in anyway—the laws have no temporal orientation. In the absence of a unified dynamical system theory of thermodynamics with Newtonian and Einsteinian mechanics, the second law is derivative to the physical laws of motion. Thus, since the (known) laws of nature are autonomous to temporal orientation, the second law implies, with identical certainty, that entropy increases both forward and backward in time from any given moment in time.
This statement, however, is not true in general; it is true only if the initial state of the universe did not begin in a highly ordered, low entropy state. However, quantum fluctuations in Higgs boson [95] particles stretched out by inflation and inflationary cosmology followed by the big bang [96] tells us that the early universe began its trajectory in a highly ordered, low entropy state, which allows us to deduce that the entropic arrow of time is not a double-headed arrow and that the future is indeed in the direction of increasing entropy. This further establishes that the concept of time flow directionality, which almost never enters in any physical theory, is a defining marvel of thermodynamics. Heat (i.e., energy), like gravity, permeates every substance in the universe and its radiation spreads to every part of spacetime. However, unlike gravity, the directional continuity of entropy and time (i.e., the entropic arrow of time) elevates thermodynamics to a sui generis physical theory of nature.

4. Modern Thermodynamics, Information Theory, and Statistical Energy Analysis

In an attempt to generalize classical thermodynamics to nonequilibrium thermodynamics, Onsager [97,98] developed reciprocity theorems for irreversible processes based on the concept of a local equilibrium that can be described in terms of state variables that are predicated on linear approximations of thermodynamic equilibrium variables. Onsager’s theorem pertains to the thermodynamics of linear systems, wherein a symmetric reciprocal relation applies between forces and fluxes. In particular, a flow or flux of matter in thermodiffusion is caused by the force exerted by the thermal gradient. Conversely, a concentration gradient causes a heat flow, an effect that has been experimentally verified for linear transport processes involving thermodiffusion, thermoelectric, and thermomagnetic effects.

Classical irreversible thermodynamics [99–101] as originally developed by Onsager characterizes the rate of entropy production of irreversible processes as a sum of the product of fluxes with their associated forces, postulating a linear relationship between the fluxes and forces. The thermodynamic fluxes in the Onsager formulation include the effects of heat conduction, flow of matter (i.e., diffusion), mechanical dissipation (i.e., viscosity), and chemical reactions. This thermodynamic theory, however, is only correct for near equilibrium processes wherein a local and linear instantaneous relation between the fluxes and forces holds.

Casimir [102] extended Onsager’s principle of macroscopic reversibility to explain the relations between irreversible processes and network theory involving the effects of electrical currents on entropy production. The Onsager-Casimir reciprocal relations treat only the irreversible aspects of system processes, and thus the theory is an algebraic theory that is primarily restricted to describing (time-independent) system steady states. In addition, the Onsager-Casimir formalism is restricted to linear systems, wherein a linearity restriction is placed on the admissible constitutive relations between the thermodynamic forces and fluxes. Another limitation of the Onsager-Casimir framework is the difficulty in providing a macroscopic description for large-scale complex dynamical systems. In addition, the Onsager-Casimir reciprocal relations are not valid on the microscopic thermodynamic level.

Building on Onsager’s classical irreversible thermodynamic theory, Prigogine [103–105] developed a thermodynamic theory of dissipative nonequilibrium structures. This theory involves kinetics describing the behavior of systems that are away from equilibrium states. However, Prigogine’s thermodynamics lacks functions of the system state, and hence, his concept of entropy for a system away from equilibrium does not have a total differential. Furthermore, Prigogine’s characterization of dissipative structures is predicated on a linear expansion of the entropy function about a particular equilibrium, and hence, is limited to the neighborhood of the equilibrium. This is a severe restriction on the applicability of this theory. In addition, his entropy cannot be calculated nor determined [106,107]. Moreover, the theory requires that locally applied exogenous heat fluxes propagate at infinite velocities across a thermodynamic body violating both experimental evidence and the principle of causality. To paraphrase Penrose, Prigogine’s thermodynamic theory at best should be regarded as a trial or deadend.

In an attempt to extend Onsager’s classical irreversible thermodynamic theory beyond a local equilibrium hypothesis, extended irreversible thermodynamics was developed in the literature [108,109]
wherein, in addition to the classical thermodynamic variables, dissipating fluxes are introduced as new independent variables providing a link between classical thermodynamics and flux dynamics. These complimentary thermodynamic variables involve nonequilibrium quantities and take the form of dissipative fluxes and include heat, viscous pressure, matter, and electric current fluxes, among others. These fluxes are associated with microscopic operators of nonequilibrium statistical mechanics and the kinetic theory of gases, and effectively describe systems with long relaxation times (e.g., low temperature solids, superfluids, and viscoelastic fluids).

Even though extended irreversible thermodynamics generalizes classical thermodynamics to nonequilibrium systems, the complementary variables are treated on the same level as the classical thermodynamic variables and hence lack any evolution equations [110–113]. To compensate for this, additional rate equations are introduced for the dissipative fluxes. Specifically, the fluxes are selected as state variables wherein the constitutive equations of Fourier, Fick, Newton, and Ohm are replaced by first-order time evolution equations that include memory and nonlocal effects.

However, unlike the classical thermodynamic variables, which satisfy conservation of mass and energy and are compatible with the second law of thermodynamics, no specific criteria are specified for the evolution equations of the dissipative fluxes. Furthermore, since every dissipative flux is formulated as a thermodynamic variable characterized by a single evolution equation with the system entropy being a function of the fluxes, extended irreversible thermodynamic theories tend to be incompatible with classical thermodynamics. Specifically, the theory yields different definitions for temperature and entropy when specialized to equilibrium thermodynamic systems.

In the last half of the twentieth century, thermodynamics was reformulated as a global nonlinear field theory with the ultimate objective to determine the independent field variables of this theory [1,114–116]. This aspect of thermodynamics, which became known as rational thermodynamics, was predicated on an entirely new axiomatic approach. As a result of this approach, modern continuum thermodynamics was developed using theories from elastic materials, viscous materials, and materials with memory [117–120]. The main difference between classical thermodynamics and rational thermodynamics can be traced back to the fact that in rational thermodynamics the second law is not interpreted as a restriction on the transformations a system can undergo, but rather as a restriction on the system’s constitutive equations.

Rational thermodynamics is formulated based on nonphysical interpretations of absolute temperature and entropy notions that are not limited to near equilibrium states. Moreover, the thermodynamic system has memory, and hence, the dynamic behavior of the system is not only determined by the present value of the thermodynamic state variables but also by the history of their past values. In addition, the second law of thermodynamics is expressed using the Clausius-Duhem inequality.

Rational thermodynamics is not a thermodynamic theory in the classical sense but rather a theory of thermomechanics of continuous media. This theory, which is also known as modern continuum thermodynamics, abandons the concept of a local equilibrium and involves general conservation laws (mass, momentum, energy) for defining a thermodynamic state of a body using a set of postulates and constitutive functionals. These postulates, which include the principles of admissibility (i.e., entropy principle), objectivity or covariance (i.e., reference frame invariance), local action (i.e., influence of a neighborhood), memory (i.e., a dynamic), and symmetry, are applied to the constitutive equations describing the thermodynamic process.

Modern continuum thermodynamics has been extended to account for nonlinear irreversible processes such as the existence of thresholds, plasticity, and hysteresis [121–124]. These extensions use convex analysis, semigroup theory, and nonlinear programming theory but can lack a clear characterization of the space over which the thermodynamical state variables evolve. The principal weakness of rational thermodynamics is that its range of applicability is limited to closed systems with a single absolute temperature. Thus, it is not applicable to condensed matter physics (e.g., diffusing mixtures or plasma). Furthermore, it does not provide a unique entropy characterization that satisfies the Clausius inequality.
More recently, a major contribution to equilibrium thermodynamics is given in [125]. This work builds on the work of Carathéodory [58,59] and Giles [126] by developing a thermodynamic system representation involving a state space on which an adiabatic accessibility relation is defined. The existence and uniqueness of an entropy function is established as a consequence of adiabatic accessibility among equilibrium states. As in Carathéodory’s work, the authors in [125] also restrict their attention to simple (possibly interconnected) systems in order to arrive at an entropy increase principle. However, it should be noted that the notion of a simple system in [125] is not equivalent to that of Carathéodory’s notion of a simple system.

Connections between thermodynamics and system theory as well as information theory have also been explored in the literature [80,127–135]. Information theory has deep connections to physics in general, and thermodynamics in particular. Many scientists have postulated that information is physical and have suggested that the bit is the irreducible kernel in the universe and it is more fundamental than matter itself, with information forming the very core of existence [136,137]. To produce change (motion) requires energy, whereas to direct this change requires information. In other words, energy takes different forms, but these forms are determined by information. Arguments about the nature of reality is deeply rooted in quantum information, which gives rise to every particle, every force field, and spacetime itself.

In quantum mechanics information can be inaccessible but not annihilated. In other words, information can never be destroyed despite the fact that imperfect system state distinguishability abounds in quantum physics, wherein the Heisenberg uncertainty principle brought the demise of determinism in the microcosm of science. The afore statement concerning the nonannihilation of information is not without its controversy in physics and is at the heart of the black hole information paradox, which resulted from the unification of quantum mechanics and general relativity.

Specifically, Hawking and Bekenstein [138,139] argued that general relativity and quantum field theory were inconsistent with the principle that information cannot be lost. In particular, as a consequence of quantum fluctuations near a black hole’s event horizon [140], they showed that black holes radiate particles, and hence, slowly evaporate. Since matter falling into a black hole carries information in its structure, organization, and quantum states, black hole evaporation via radiation obliterates information.

However, using Richard Feynman’s sum over histories path integral formulation of quantum theory to the topology of spacetime [141], Hawking later showed that quantum gravity is unitary (i.e., the sum of probabilities for all possible outcomes of a given event is unity) and that black holes are never unambiguously black. That is, black holes slowly dissipate before they ever truly form allowing radiation to contain information, and hence, information is not lost, obviating the information paradox.

In quantum mechanics the Heisenberg uncertainty principle is a consequence of the fact that the outcome of an experiment is affected, or even determined, when observed. The Heisenberg uncertainty principle states that it is impossible to measure both the position and momentum of a particle with absolute precision at a microscopic level, and the product of the uncertainties in these measured values is of the order of magnitude of the Planck constant. The determination of energy and time is also subject to the same uncertainty principle. The principle is not a statement about our inability to develop accurate measuring instruments, but rather a statement about an intrinsic property of nature; namely, nature has an inherent indeterminacy. This is a consequence of the fact that any attempt at observing nature will disturb the system under observation resulting in a lack of precision.

Quantum mechanics provides a probabilistic theory of nature, wherein the equations describe the average behavior of a large collection of identical particles and not the behavior of individual particles. Einstein maintained that the theory was incomplete albeit a good approximation in describing nature. He further asserted that when quantum mechanics had been completed, it would deal with certainties. In a letter to Max Born he states his famous God does not play dice dictum writing [142] (p. 90): “The theory produces a great deal but hardly brings us closer to the secret of the Old One. I am at all
events convinced that He does not play dice”. A profound ramification of the Heisenberg uncertainty principle is that the macroscopic principle of causality does not apply at the atomic level.

Information theory addresses the quantification, storage, and communication of information. The study of the effectiveness of communication channels in transmitting information was pioneered by Shannon [143]. Information is encoded, stored (by codes), transmitted through channels of limited capacity, and then decoded. The effectiveness of this process is measured by the Shannon capacity of the channel and involves the entropy of a set of events that measure the uncertainty of this set. These channels function as input-output devices that take letters from an input alphabet and transmit letters to an output alphabet with various error probabilities that depend on noise. Hence, entropy in an information-theoretic context is a measure of information uncertainty. Simply put—information is not free and is linked to the cost of computing the behavior of matter and energy in our universe [144]. For an excellent exposition of these different facets of thermodynamics see [145].

Thermodynamic principles have also been repeatedly used in coupled mechanical systems to arrive at energy flow models. Specifically, in an attempt to approximate high-dimensional dynamics of large-scale structural (oscillatory) systems with a low-dimensional diffusive (non-oscillatory) dynamical model, structural dynamicists have developed thermodynamic energy flow models using stochastic energy flow techniques. In particular, statistical energy analysis (SEA) predicated on averaging system states over the statistics of the uncertain system parameters has been extensively developed for mechanical and acoustic vibration problems [81,146–150]. The aim of SEA is to establish that many concepts of energy flow modeling in high-dimensional mechanical systems have clear connections with statistical mechanics of many particle systems, and hence, the second law of thermodynamics applies to large-scale coupled mechanical systems with modal energies playing the role of temperatures.

Thermodynamic models are derived from large-scale dynamical systems of discrete subsystems involving stored energy flow among subsystems based on the assumption of weak subsystem coupling or identical subsystems. However, the ability of SEA to predict the dynamic behavior of a complex large-scale dynamical system in terms of pairwise subsystem interactions is severely limited by the coupling strength of the remaining subsystems on the subsystem pair. Hence, it is not surprising that SEA energy flow predictions for large-scale systems with strong coupling can be erroneous. From the rigorous perspective of dynamical systems theory, the theoretical foundations of SEA remain inadequate since well-defined mathematical assumptions of the theory are not adequately delineated.

Alternatively, a deterministic thermodynamically motivated energy flow modeling for structural systems is addressed in [151–153]. This approach exploits energy flow models in terms of thermodynamic energy (i.e., the ability to dissipate heat) as opposed to stored energy and is not limited to weak subsystem coupling. A stochastic energy flow compartmental model (i.e., a model characterized by energy conservation laws) predicated on averaging system states over the statistics of stochastic system exogenous disturbances is developed in [130]. The basic result demonstrates how linear compartmental models arise from second-moment analysis of state space systems under the assumption of weak coupling. Even though these results can be potentially applicable to linear large-scale dynamical systems with weak coupling, such connections are not explored in [130]. With the notable exception of [150], and more recently [111–113,154], none of the aforementioned SEA-related works addresses the second law of thermodynamics involving entropy notions in the energy flow between subsystems.

Motivated by the manifestation of emergent behavior of macroscopic energy transfer in crystalline solids modeled as a lattice of identical molecules involving undamped vibrations, the authors in [155] analyze energy equipartition in linear Hamiltonian systems using average-preserving symmetries. Specifically, the authors consider a Lie group of phase space symmetries of a linear Hamiltonian system and characterize the subgroup of symmetries whose elements are also symmetries of every Hamiltonian system and preserve the time averages of quadratic Hamiltonian functions along system trajectories. In the very specific case of distinct natural frequencies and a two-degree-of-freedom system consisting
of an interconnected pair of identical undamped oscillators, the authors show that the time-averaged oscillator energies reach an equipartitioned state. For this limited case, this result shows that time averaging leads to the emergence of damping in lossless Hamiltonian dynamical systems.

5. Dynamical Systems

Dynamical systems theory provides a universal mathematical formalism predicated on modern analysis and has become the prevailing language of modern science as it provides the foundation for unlocking many of the mysteries in nature and the universe that involve spatial and temporal evolution. Given that irreversible thermodynamic systems involve a definite direction of evolution, it is natural to merge the two universalisms of thermodynamics and dynamical systems under a single compendium, with the latter providing an ideal language for the former.

A system is a combination of components or parts that is perceived as a single entity. The parts making up the system may be clearly or vaguely defined. These parts are related to each other through a particular set of variables, called the states of the system, that, together with the knowledge of any system inputs, completely determine the behavior of the system at any given time. A dynamical system is a system whose state changes with time. Dynamical system theory was fathered by Henri Poincaré [156–158], sturdily developed by Birkhoff [159,160], and has evolved to become one of the most universal mathematical formalisms used to explain system manifestations of nature that involve time.

A dynamical system can be regarded as a mathematical model structure involving an input, state, and output that can capture the dynamical description of a given class of physical systems. Specifically, a closed dynamical system consists of three elements—namely, a setting called the state space which is assumed to be Hausdorff and in which the dynamical behavior takes place, such as a torus, topological space, manifold, or locally compact metric space; a mathematical rule or dynamic which specifies the evolution of the system over time; and an initial condition or state from which the system starts at some initial time.

An open dynamical system interacts with the environment through system inputs and system outputs and can be viewed as a precise mathematical object which maps exogenous inputs (causes, disturbances) into outputs (effects, responses) via a set of internal variables, the state, which characterizes the influence of past inputs. For dynamical systems described by ordinary differential equations, the independent variable is time, whereas spatially distributed systems described by partial differential equations involve multiple independent variables reflecting, for example, time and space.

The state of a dynamical system can be regarded as an information storage or memory of past system events. The set of (internal) states of a dynamical system must be sufficiently rich to completely determine the behavior of the system for any future time. Hence, the state of a dynamical system at a given time is uniquely determined by the state of the system at the initial time and the present input to the system. In other words, the state of a dynamical system in general depends on both the present input to the system and the past history of the system. Even though it is often assumed that the state of a dynamical system is the least set of state variables needed to completely predict the effect of the past upon the future of the system, this is often a convenient simplifying assumption.

Ever since its inception, the basic questions concerning dynamical system theory have involved qualitative solutions for the properties of a dynamical system; questions such as: For a particular initial system state, Does the dynamical system have at least one solution? What are the asymptotic properties of the system solutions? How are the system solutions dependent on the system initial conditions? How are the system solutions dependent on the form of the mathematical description of the dynamic of the system? How do system solutions depend on system parameters? And how do system solutions depend on the properties of the state space on which the system is defined?

Determining the rule or dynamic that defines the state of physical systems at a given future time from a given present state is one of the central problems of science. Once the flow or dynamic of a dynamical system describing the motion of the system starting from a given initial state is given,
dynamical system theory can be used to describe the behavior of the system states over time for different initial conditions. Throughout the centuries—from the great cosmic theorists of ancient Greece [161] to the present-day quest for a unified field theory—the most important dynamical system is our universe. By using abstract mathematical models and attaching them to the physical world, astronomers, mathematicians, and physicists have used abstract thought to deduce something that is true about the natural system of the cosmos.

The quest by scientists, such as Brahe, Kepler, Galileo, Newton, Huygens, Euler, Lagrange, Laplace, and Maxwell, to understand the regularities inherent in the distances of the planets from the Sun and their periods and velocities of revolution around the Sun led to the science of dynamical systems as a branch of mathematical physics. Isaac Newton, however, was the first to model the motion of physical systems with differential equations. Newton’s greatest achievement was the rediscovery that the motion of the planets and moons of the solar system resulted from a single fundamental source—the gravitational attraction of the heavenly bodies. This discovery dates back to Aristarkhos’ (310–230 B.C.) heliocentric theory of planetary motion and Hipparkhos’ (190–120 B.C.) dynamical theory of planetary motions predicated on planetary attractions toward the Sun by a force that is inversely proportional to the square of the distance between the planets and the Sun [162] (p. 304).

Many of the concepts of Newtonian mechanics, including relative motion, centrifugal and centripetal force, inertia, projectile motion, resistance, gravity, and the inverse-square law were known to the Hellenistic scientists [162]. For example, Hipparkhos’ work *On bodies thrusting down because of gravity* (Περὶ τῶν διὰ βαρότητα κάτω φερομένων) clearly and correctly describes the effects of gravity on projectile motion. In Ploutarkhos’ (46–120 A.D.) work *De facie quae in orbe lunae apparet* (On the light glowing on the Moon) he clearly describes the notion of gravitational interaction between heavenly bodies stating that [163] “just as the sun attracts to itself the parts of which it consists, so does the earth . . .” (p. 304).

Newton himself wrote in his *Classical Scholia* [162] (p. 376): “Pythagoras . . . applied to the heavens the proportions found through these experiments [on the pitch of sounds made by weighted strings], and learned from that the harmonies of the spheres. So, by comparing those weights with the weights of the planets, and the intervals in sound with the intervals of the spheres, and the lengths of string with the distances of the planets [measured] from the center, he understood through the heavenly harmonies that the weights of the planets toward the sun . . . are inversely proportional to the squares of their distances”. This admittance of the prior knowledge of the inverse square law predates Hooke’s thoughts of explaining Kepler’s laws out of the inverse square law communicated in a letter on 6 January 1680 to Newton by over two millennia.

It is important to stress here that what are erroneously called *Newton’s laws of motion* in the literature were first discovered by Kepler, Galileo, and Descartes, with the latter first stating the law of inertia in its modern form. Namely, when viewed in an inertial reference frame, a body remains in the same state unless acted upon by a net force; and unconstrained motion follows a rectilinear path. Newton and Leibnitz independently advanced the basic dynamical tool invented two millennia earlier by Archimedes—the calculus [164,165] with Euler being the first one to explicitly write down the second law of motion as an equation involving an applied force acting on a body being equal to the time rate of change of its momentum. Newton, however, deduced a physical hypothesis—the law of universal gravitation involving an inverse-square law force—in precise mathematical form deriving (at the time) a cosmic dynamic using Euclidian geometry and not differential calculus (i.e., differential equations).

In his magnum opus *Philosophiae Naturalis Principia Mathematica* [166] Newton investigated whether a small perturbation would make a particle moving in a plane around a center of attraction continue to move near the circle, or diverge from it. Newton used his analysis to analyze the motion of the moon orbiting the Earth. Numerous astronomers and mathematicians who followed made significant contributions to dynamical system theory in an effort to show that the observed deviations of planets and satellites from fixed elliptical orbits were in agreement with Newton’s principle of universal gravitation. Notable contributions include the work of Torricelli [167], Euler [168],
Newtonian mechanics developed into the first field of modern science—dynamical systems as a branch of mathematical physics—wherein the circular, elliptical, and parabolic orbits of the heavenly bodies of our solar system were no longer fundamental determinants of motion, but rather approximations of the universal laws of the cosmos specified by governing differential equations of motion. In the past century, dynamical system theory has become one of the most fundamental fields of modern science as it provides the foundation for unlocking many of the mysteries in nature and the universe that involve the evolution of time. Dynamical system theory is used to study ecological systems, geological systems, biological systems, economic systems, neural systems, and physical systems (e.g., mechanics, fluids, magnetic fields, galaxies, etc.), to cite but a few examples. The long absence of a dynamical system formalism in classical thermodynamics, and physics in general, is quite disturbing and in our view largely responsible for the monomeric state of classical thermodynamics.

6. System Thermodynamics: A Postmodern Approach

In contrast to mechanics, which is based on a dynamical system theory, classical thermodynamics (i.e., thermostatics) is a physical theory and does not possess equations of motion. Moreover, very little work has been done in obtaining extensions of thermodynamics for systems out of equilibrium. These extensions are commonly known as thermodynamics of irreversible processes or modern irreversible thermodynamics in the literature [103,178]. Such systems are driven by the continuous flow of matter and energy, are far from equilibrium, and often develop into a multitude of states. Connections between local thermodynamic subsystem interactions of these systems and the globally complex thermo dynamical system behavior is often elusive. This statement is true for nature in general and was most eloquently stated first by Herakleitos in his 123rd fragment—Φύσις κρύπτεσθαι φιλεί (Nature loves to hide).

These complex thermodynamic systems involve spatio-temporally evolving structures and can exhibit a hierarchy of emergent system properties. These systems are known as dissipative systems [3,9] and consume energy and matter while maintaining their stable structure by dissipating entropy to the environment. All living systems are dissipative systems, the converse, however, is not necessarily true. Dissipative living systems involve pattern interactions by which life emerges. This nonlinear interaction between the subsystems making up a living system is characterized by autopoiesis (self-creation). In the physical universe, billions of stars and galaxies interact to form self-organizing dissipative nonequilibrium structures [94,179]. The fundamental common phenomenon among nonequilibrium (i.e., dynamical) systems are that they evolve in accordance to the laws of (nonequilibrium) thermodynamics.

Building on the work of nonequilibrium thermodynamic structures [103,105], Sekimoto [180–184] introduced a stochastic thermodynamic framework predicated on Langevin dynamics in which fluctuation forces are described by Brownian motion. In this framework, the classical thermodynamic notions of heat, work, and entropy production are extended to the level of individual system trajectories of nonequilibrium ensembles. Specifically, system state trajectories are sample continuous and are characterized by a Langevin equation for each individual sample path and a Fokker-Planck equation for the entire ensemble of trajectories. For such systems, energy conservation holds along fluctuating trajectories of the stochastic Markov process and the second law of thermodynamics is obtained as an ensemble property of the process. In particular, various fluctuation theorems [184–195] are derived that constrain the probability distributions for the exchanged heat, mechanical work, and entropy production depending on the nature of the stochastic Langevin system dynamics.

Even though stochastic thermodynamics is applicable to a single realization of the Markov process under consideration with the first and second laws of thermodynamics holding for nonequilibrium systems, the framework only applies to multiple time-scale systems with a few observable slow degrees
of freedom. The unobservable degrees of freedom are assumed to be fast, and hence, always constrained to the equilibrium manifold imposed by the instantaneous values of the observed slow degrees of freedom. Furthermore, if some of the slow variables are not accessible, then the system dynamics are no longer Markovian. In this case, defining a system entropy is virtually impossible. In addition, it is unclear whether fluctuation theorems expressing symmetries of the probability distribution functions for thermodynamic quantities can be grouped into universal classes characterized by asymptotics of these distributions. Moreover, it is also unclear whether there exist system variables that satisfy the transitive equilibration property of the zeroth law of thermodynamics for nonequilibrium stochastic thermodynamic systems.

In an attempt to create a generalized theory of evolution mechanics by unifying classical mechanics with thermodynamics, the authors in [111–113,196] developed a new and novel framework of dynamical thermodynamics based on the concept of tribo-fatigue entropy. This framework, known as damage mechanics [111,196] or mechanothermodynamics [112,113], involves an irreversible entropy function along with its generation rate that captures and quantifies system aging. Specifically, the second law is formulated analytically for organic and inorganic bodies, and the system entropy is determined by a damageability process predicated on mechanical and thermodynamic effects resulting in system state changes.

In [3,9], the authors develop a postmodern framework for thermodynamics that involves open interconnected dynamical systems that exchange matter and energy with their environment in accordance with the first law (conservation of energy) and the second law (nonconservation of entropy) of thermodynamics. Specifically, we use the state space formalism to construct a mathematical model that is consistent with basic thermodynamic principles. This is in sharp contrast to classical thermodynamics wherein an input-output description of the system is used. However, it is felt that a state space formulation is essential for developing a thermodynamic model with enough detail for describing the thermal behavior of heat and dynamical energy. In addition, such a model is crucial in accounting for internal system properties captured by compartmental system dynamics characterizing conservation laws, wherein subsystem energies can only be transported, stored, or dissipated but not created.

If a physical system possesses these conservation properties externally, then there exists a high possibility that the system possesses these properties internally. Specifically, if the system possesses conservation properties internally and does not violate any input-output behavior of the physical system established by experimental evidence, then the state space model is theoretically credible. The governing physical laws of nature impose specific relations between the system state variables, with system interconnections captured by shared variables among subsystems. An input-state-output model is ideal in capturing system interconnections with the environment involving the exchange of matter, energy, or information, with inputs serving to capture the influence of the environment on the system, outputs serving to capture the influence of the system on the environment, and internal feedback interconnections—via specific output-to-input assignments—capturing interactions between subsystems.

The state space model additionally enforces the fundamental property of causality, that is, nonanticipativity of the system, wherein future values of the system input do not influence past and present values of the system output. More specifically, the system state, and hence, the system output before a certain time are not affected by the values of the input after that time; a principal holding for all physical systems verified by experiment.

Another important notion of state space modeling is that of system realization involving the construction of the state space, the system dynamic, and the system output that yields a dynamical system in state space form and generates the given input-output map—established by experiment—through a suitable choice of the initial system state. This problem has been extensively addressed in dynamical systems theory and leads to the fact that every continuous input-output dynamical system has a continuous realization in state space form; for details see [197–199].
In a dynamical systems framework of thermodynamics, symmetry can spontaneously occur by invoking the two fundamental axioms of the science of heat. Namely, (i) if the energies in the connected subsystems of an interconnected system are equal, then energy exchange between these subsystems is not possible, and (ii) energy flows from more energetic subsystems to less energetic subsystems. These axioms establish the existence of a global system entropy function as well as equipartition of energy [3] in system thermodynamics; an emergent behavior in thermodynamic systems. Hence, in complex interconnected thermodynamic systems, higher symmetry is not a property of the system’s parts but rather emerges as a result of the nonlinear subsystem interactions.

The goal of the recent monograph [9] is directed towards building on the results of [3] to place thermodynamics on a system-theoretic foundation by combining the two universalisms of thermodynamics and dynamical systems theory under a single umbrella so as to harmonize it with classical mechanics. In particular, the author develops a novel formulation of thermodynamics that can be viewed as a moderate-sized system theory as compared to statistical thermodynamics. This middle-ground theory involves large-scale dynamical system models characterized by ordinary deterministic and stochastic differential equations, as well as infinite-dimensional models characterized by partial differential equations that bridge the gap between classical and statistical thermodynamics.

Specifically, since thermodynamic models are concerned with energy flow among subsystems, we use a state space formulation to develop a nonlinear compartmental dynamical system model that is characterized by energy conservation laws capturing the exchange of energy and matter between coupled macroscopic subsystems. Furthermore, using graph-theoretic notions, we state two thermodynamic axioms consistent with the zeroth and second laws of thermodynamics, which ensure that our large-scale dynamical system model gives rise to a thermodynamically consistent energy flow model. Specifically, using a large-scale dynamical systems theory perspective for thermodynamics, we show that our compartmental dynamical system model leads to a precise formulation of the equivalence between work energy and heat in a large-scale dynamical system.

Since our thermodynamic formulation is based on a large-scale dynamical system theory involving the exchange of energy with conservation laws describing transfer, accumulation, and dissipation between subsystems and the environment, our framework goes beyond classical thermodynamics characterized by a purely empirical theory, wherein a physical system is viewed as an input-output black box system. Furthermore, unlike classical thermodynamics, which is limited to the description of systems in equilibrium states, our approach addresses nonequilibrium thermodynamic systems. This allows us to connect and unify the behavior of heat as described by the equations of thermal transfer and as described by classical thermodynamics. This exposition further demonstrates that these disciplines of classical physics are derivable from the same principles and are part of the same scientific and mathematical framework.

Our nonequilibrium thermodynamic framework goes beyond classical irreversible thermodynamics developed by Onsager [97,98] and further extended by Casimir [102] and Prigogine [103–105], which, as discussed in Section 4, fall short of a complete dynamical theory. Specifically, their theories postulate that the local instantaneous thermodynamic variables of the system are the same as that of the system in equilibrium. This implies that the system entropy in a neighborhood of an equilibrium is dependent on the same variables as those at equilibrium violating Gibb’s principle. In contrast, the proposed system thermodynamic formalism brings classical thermodynamics within the framework of modern nonlinear dynamical systems theory, thus providing information about the dynamical behavior of the thermodynamic state variables between the initial and final equilibrium system states.

Next, we give a deterministic definition of entropy for a large-scale dynamical system that is consistent with the classical thermodynamic definition of entropy, and we show that it satisfies a Clausius-type inequality leading to the law of entropy nonconservation. However, unlike classical thermodynamics, wherein entropy is not defined for arbitrary states out of equilibrium, our definition of entropy holds for nonequilibrium dynamical systems. Furthermore, we introduce a new and dual notion to entropy—namely, ectropy [200]—as a measure of the tendency of a large-scale dynamical
system to do useful work and grow more organized, and we show that conservation of energy in an adiabatically isolated thermodynamically consistent system necessarily leads to nonconservation of entropy and entropy. Hence, for every dynamical transformation in an adiabatically isolated thermodynamically consistent system, the entropy of the final system state is greater than or equal to the entropy of the initial system state.

Then, using the system entropy as a Lyapunov function candidate, we show that in the absence of energy exchange with the environment our thermodynamically consistent large-scale nonlinear dynamical system model possesses a continuum of equilibria and is semistable, that is, it has convergent subsystem energies to Lyapunov stable energy equilibria determined by the large-scale system initial subsystem energies. In addition, we show that the steady-state distribution of the large-scale system energies is uniform, leading to system energy equipartitioning corresponding to a minimum entropy and a maximum entropy equilibrium state.

For our thermodynamically consistent dynamical system model, we further establish the existence of a unique continuously differentiable global entropy and entropy function for all equilibrium and nonequilibrium states. Using these global entropy and entropy functions, we go on to establish a clear connection between thermodynamics and the arrow of time. Specifically, we rigorously show a state irrecoverability and hence a state irreversibility [70,201] nature of thermodynamics. In particular, we show that for every nonequilibrium system state and corresponding system trajectory of our thermodynamically consistent large-scale nonlinear dynamical system, there does not exist a state such that the corresponding system trajectory completely recovers the initial system state of the dynamical system and at the same time restores the energy supplied by the environment back to its original condition.

This, along with the existence of a global strictly increasing entropy function on every nontrivial system trajectory, gives a clear time-reversal asymmetry characterization of thermodynamics, establishing an emergence of the direction of time flow. In the case where the subsystem energies are proportional to subsystem temperatures, we show that our dynamical system model leads to temperature equipartition, wherein all the system energy is transferred into heat at a uniform temperature. Furthermore, we show that our system-theoretic definition of entropy and the newly proposed notion of entropy are consistent with Boltzmann’s kinetic theory of gases involving an \(n\)-body theory of ideal gases divided by diathermal walls. Finally, these results are generalized to continuum thermodynamics involving infinite-dimensional energy flow conservation models. This dynamical systems theory of thermodynamics can potentially provide deeper insights into some of the most perplexing questions about the origins and fabric of our universe that require dynamical system models that are far from equilibrium.

7. Thermodynamics of Living Systems

As noted in the Introduction, the universal irreversibility associated with the second law of thermodynamics responsible for the enfeeblement and eventual demise of the universe has led numerous writers, historians, philosophers, and theologians to ask the momentous question: How is it possible for life to come into being in a universe governed by a supreme law that impedes the very existence of life? This question, followed by numerous erroneous apodoses that life defies the second law and that the creation of life is an unnatural state in the constant ravaging progression of diminishing order in the universe, has been used by creationists, evolution theorists, and intelligent designers to promote their own points of view regarding the existence or nonexistence of a supreme being. However, what many bloviators have failed to understand is that these assertions on the origins and subsistence of life along with its evolution present merely opinions of simple citizens and amateur scientists; they have very little to do with a rigorous understanding of the supreme law of nature—the second law of thermodynamics.

Life and evolution do not violate the second law of thermodynamics; we know of nothing in Nature that violates this absolute law. In fact, the second law of thermodynamics is not an
impediment to understanding life but rather is a necessary condition for providing a complete elucidation of living processes. The nonequilibrium state of the universe is essential in extracting free energy [202] and destroying [203] this energy to create entropy and consequently maintaining highly organized structures in living systems. The entire fabric of life on Earth necessitates an intricate balance of organization involving a highly ordered, low entropy state. This in turn requires dynamic interconnected structures governed by biological and chemical principles, including the principle of natural selection, which do not violate the thermodynamic laws as applied to the whole system in question.

Organization here refers to a *large-scale system* with multiple interdependent parts forming ordered (possibly organic) spatio-temporal unity, and wherein the system is greater than the sum of its parts. Order here refers to arrangements involving patterns. Both organization and order are ubiquitous in nature. For example, biology has shown us that many species of animals such as insect swarms, ungulate flocks, fish schools, ant colonies, and bacterial colonies self-organize in nature.

These biological aggregations give rise to remarkably complex global behaviors from simple local interactions between a large number of relatively unintelligent agents without the need for a centralized architecture. The spontaneous development (i.e., self-organization) of these autonomous biological systems and their spatio-temporal evolution to more complex states often appears without any external system interaction. In other words, structural morphing into coherent groups is internal to the system and results from local interactions among subsystem components that are either dependent or independent of the physical nature of the individual components. These local interactions often comprise a simple set of rules that lead to remarkably complex and robust behaviors. Robustness here refers to insensitivity of individual subsystem failures and unplanned behavior at the individual subsystem level.

The second law does not preclude subsystem parts of a large-scale system from achieving a more ordered state in time, and hence, evolving into a more organized and complex form. The second law asserts that the *total system entropy of an isolated system* always increases. Thus, even though the creation of a certain measure of life and order on Earth involves a dynamic transition to a more structured and organized state, it is unavoidably accompanied by an even greater measure of death and disorder in the (almost) isolated system encompassing the Sun and the Earth. In other words, although some parts of this isolated system will gain entropy (d$S > 0$) and other parts will lose entropy (d$S < 0$), the total entropy of this isolated system will always increase over time.

Our Sun [204] is the predominant energy source that sustains life on Earth in organisms composed of mutually interconnected parts maintaining various vital processes. These organisms maintain a highly organized state by extracting free energy from their surroundings, subsumed within a larger system, and processing this energy to maintain a low-entropy state. In other words, living systems continuously extract free energy from their environment and export waste to the environment thereby producing entropy. Thus, the creation of internal system organization and complexity (negative entropy change) through energy dispersal (i.e., elimination of energy gradients) necessary for sustaining life is balanced by a greater production of entropy externally acquiescing to the second law.

The spontaneous dispersal of energy gradients by every complex animate or inanimate mechanism leads to the law of entropy increase; that is, in an isolated system, entropy is generated and released to the environment resulting in an overall system entropy increase. However, Janes’ principle of maximum entropy production [205,206], which states that among all possible energy dissipation paths that a system may move through spontaneously, the system will select the path that will maximize its entropy production rate, adds yet another level of subtlety to all dissipative structures in the universe including all forms of life. In particular, since complex organized structures efficiently dissipate energy gradients, Nature, whenever possible, optimally transitions (in the fastest possible way) these complex structures to dissipate energy gradients spontaneously creating order, and thus, maximizing entropy production. This results in a universe that exists at the *expense* of life and order—the enthroning emblem of the second law of thermodynamics.
The degradation of free energy sources by living organisms for maintaining nonequilibrium configurations as well as their local level of organization at the expense of a larger global entropy production was originally recognized by Schrödinger [207]. Schrödinger maintained that life was comprised from two fundamental processes in nature; namely, order from order and order from disorder. He postulated that life cannot exist without both processes; and that order from disorder is necessary to generate life, whereas order from order is necessary to ensure the propagation of life.

In 1953 Watson and Crick [208] addressed the first process by observing that genes generate order from order in a species; that is, inherited traits are encoded and passed down from parents to offspring within the DNA double helix. The dynamical systems framework of thermodynamics can address the fundamental process of the order from disorder mystery by connecting system biology to the second law of thermodynamics. More specifically, a dynamical systems theory of thermodynamics can provide an understanding of how the microscopic molecular properties of living systems lead to the observed macroscopic properties of these systems.

The external source of energy that sustains life on Earth—the Sun—is of course strictured by the second law of thermodynamics. However, it is important to recognize that this energy source is a low-entropy energy source and that the energy flow over a twenty-four hour (day-night) cycle from the Sun to the Earth and thenceforth the Earth to the darkness of space is essentially balanced. This balance is modulo the small amounts of heat generated by the burning of fossil fuels and global warming, as well as the radioactive material from volcanic deep ocean vent ecosystems. This balance can also be effected by Milankovich cycles (i.e., changes in the Earth’s climate due to variations in eccentricity, axial tilt, and procession of the Earth’s orbit) as well as variations in solar luminosity. If this intricate energy balance did not take place, then the Earth would accelerate towards a thermal equilibrium and all life on Earth would cease to exist.

The subsistence of life on Earth is due to the fact that the Sun’s temperature is far greater than that of dark space, and hence, yellow light photons [209] emitted from the Sun’s atoms have a much higher frequency than the infrared photons the Earth dissipates into dark space. This energy exchange includes reflection and scattering by the oceans, polar ice caps, barren terrain, and clouds; atmosphere absorption leading to energy dissipation through thermal infrared emissions and chemical pathways; ocean absorption via physicochemical and biological processes; passive absorption and reemission of infrared photons by barren terrain; latent heat generation through hydrological cycles; and the thermal dissipation from the Earth’s heat sources.

Since the photons absorbed by the Earth from the Sun have a much higher frequency, and hence, shorter wavelength than those dissipated to space by the Earth, it follows from Planck’s quantum formula [210] relating the energy of each photon of light to the frequency of the radiation it emits that the average energy flow into the Earth from the Sun by each individual photon is far greater than the dissipated energy by the Earth to dark space. This implies that the number of different photon microstates carrying the energy from the Earth to the darkness of space is far greater than the number of microstates carrying the same energy from the Sun to the Earth. Thus, it follows from Boltzmann’s entropy formula that the incident energy from the Sun is at a considerably lower entropy level than the entropy produced by the Earth through energy dissipation.

It is important to stress that the Sun continuously supplies the Earth with high-grade free energy (i.e., low-entropy) that keeps the Earth’s entropy low. The Earth’s low entropy is predominately maintained through phototrophs (i.e., plants, algae, cyanobacteria, and all photosynthesizers) and chemotrophs (e.g., iron and manganese oxidizing bacteria found in igneous lava rock) which get their free energy from solar photons and inorganic compounds. For example, leaves on most green plants have a relatively large surface area per unit volume allowing them to capture high-grade free energy from the Sun. Specifically, photon absorption by photosynthesizers generate pathways for energy conversion coupled to photochemical, chemical, and electrochemical reactions leading to high-grade energy. This energy is converted into useful work to maintain the plants in a complex, high-energy organized state via photosynthesis supporting all heterotrophic life on Earth.
Energy dissipation leading to the production of entropy is a direct consequence of the fact that the Sun’s temperature ($T_{\text{sun}} = 5760$ K) is at a much higher temperature than that of the Earth’s temperature ($T_{\text{earth}} = 255$ K) and, in turn, the Earth’s temperature is higher than the temperature of outer space. Even though the energy coming to the Earth from the Sun is balanced by the energy radiated by the Earth into dark space, the photons reaching the Earth’s surface from the Sun have a higher energy content (i.e., shorter wavelength) than the photons radiated by the Earth into the dark sky.

In particular, for every one solar photon infused by the Earth, the Earth diffuses approximately twenty photons into the dark sky with approximately twenty times longer wavelengths [211,212]. Moreover, since the entropy of photons is proportional to the number of photons [212], for every solar photon the Earth absorbs at low entropy, the Earth dilutes the energy of the solar photon among twenty photons and radiates them into deep space. In other words, the Earth is continuously extracting high-grade free energy from the Sun and exporting degraded energy, thereby producing entropy and maintaining low entropy structures on Earth.

The same conclusion can also be arrived at using a macroscopic formulation of entropy. In particular, following a similar analysis as in [212], note that the energy reaching the Earth $dQ$ is equal to the energy radiated by the Earth $-dQ$, $T_{\text{earth}} < T_{\text{sun}}$, and

$$
\frac{dS_{\text{in}}}{T_{\text{sun}}} \geq \frac{dQ}{T_{\text{sun}}} \quad (1)
$$

$$
\frac{dS_{\text{out}}}{T_{\text{earth}}} \geq -\frac{dQ}{T_{\text{earth}}} \quad (2)
$$

where $dS_{\text{in}}$ and $dS_{\text{out}}$ is the change in entropy from the absorption and emission of photons, respectively.

Now, under steady Earth conditions, that is, constant average temperature, constant internal energy $U_{\text{earth}}$, and constant entropy $S_{\text{earth}}$, the change in entropy of the Earth is

$$
dS_{\text{earth}} = dS_{\text{photons}} + dS_{\text{dissip}} = 0, \quad (3)
$$

where $dS_{\text{photons}} = dS_{\text{in}} + dS_{\text{out}}$ and $dS_{\text{dissip}}$ denotes the entropy produced by all the Earth’s dissipative structures. Hence,

$$
dS_{\text{photons}} = -dS_{\text{dissip}} < 0. \quad (4)
$$

Thus, $|dS_{\text{in}}| < |dS_{\text{out}}|$, and hence, it follows from (1) and (2) that

$$
\left| \frac{dS_{\text{out}}}{dS_{\text{in}}} \right| \leq \frac{T_{\text{sun}}}{T_{\text{earth}}} = \frac{5760}{255} \approx 20, \quad (5)
$$

which implies that the Earth produces at most twenty times as much entropy as it receives from the Sun. Thus, the net change in the Earth’s entropy from solar photon infusion and infrared photon radiation is given by

$$
dS_{\text{photons}} = dS_{\text{in}} + dS_{\text{out}} \geq dQ \left( \frac{1}{T_{\text{sun}}} - \frac{1}{T_{\text{earth}}} \right)
= -\frac{dQ}{T_{\text{earth}}} \left( 1 - \frac{T_{\text{earth}}}{T_{\text{sun}}} \right)
= -0.95 \frac{dQ}{T_{\text{earth}}}. \quad (6)
$$

It is important to note here that even through there is a net decrease in the entropy of the Earth from the absorption and emission of photons, this is balanced by the increase in entropy exported from all of the low entropy dissipative structures on Earth. Hence, the exported waste by living systems
and dissipative structures into the environment results in an entropy production that does not lower the entropy of the Earth, but rather sustains the Earth’s entropy at a constant low level. To maintain this low level entropy, a constant supply of free energy is needed.

Specifically, since the amount of the Earth’s free energy is not increasing under steady Earth conditions it follows that

$$\text{d}F_{\text{earth}} = \text{d}F_{\text{photons}} + \text{d}F_{\text{dissp}} = 0,$$

where \( \text{d}F_{\text{photons}} \) denotes the amount of free energy received by the Earth through solar photons and \( \text{d}F_{\text{dissp}} \) is the amount of free energy dissipated by the Earth’s dissipative structures (e.g., a myriad of life forms, cyclones, thermohaline circulation, storms, atmospheric circulation, convection cells, etc.), and hence, \( \text{d}F_{\text{photons}} = -\text{d}F_{\text{dissp}} \). Finally, noting that

$$\text{d}F_{\text{earth}} = \text{d}U_{\text{earth}} - T_{\text{earth}} \text{d}S_{\text{earth}} - S_{\text{earth}} \text{d}T_{\text{earth}},$$

and assuming steady Earth conditions, it follows from (8) that

$$\text{d}F_{\text{photons}} + \text{d}F_{\text{dissp}} = -T_{\text{earth}} (\text{d}S_{\text{photons}} + \text{d}S_{\text{dissp}}).$$

Now, equating analogous contributing terms in (9) gives \( \text{d}F_{\text{photons}} = -T_{\text{earth}} \text{d}S_{\text{photons}}, \) which using (6), yields

$$\text{d}F_{\text{photons}} \leq 0.95dQ.$$

Thus, it follows (10) that, under steady conditions, at most 95% of the Sun’s energy can be used to perform useful work on Earth.

8. Thermodynamics and the Origin of Life

In Section 7, we provided a rationalization for the existence of living systems using thermodynamics. The origins of life, however, is a far more complex matter. At the heart of this complexity lies the interplay of Schrödinger’s two fundamental processes for life; namely, order from order and order from disorder [207]. These fundamental processes involve complex self-organizing living organisms that exhibit remarkable robustness over their healthy stages of life as well as across generations, through heredity.

The order from disorder robustness in maintaining a living system requires that the free energy flow through the system is such that the negative (local) entropy production rate generated by the living system \( \text{d}S_{\text{ext}} \) is greater than the internal entropy rate \( \text{d}S_{\text{int}} \) accumulated by the living system from irreversible processes taking place within the system, that is, \( |\text{d}S_{\text{ext}}| > \text{d}S_{\text{int}} > 0 \). Alternatively, the order from order robustness is embedded within the genes of the living organism encapsulating information, wherein every cell transmits and receives as well as codes and decodes life’s blueprint passed down through eons.

Deoxyribonucleic acid (DNA) is the quintessential information molecule that carries the genetic instructions of all known living organisms. The human DNA is a cellular information processor composed of an alphabet code sequence of six billion bits. DNA, Ribonucleic acid (RNA), proteins, and polysaccharides (i.e., complex carbohydrates) are key types of macromolecules that are essential for all forms of life. The macromolecules (DNA, RNA, proteins, and polysaccharides) of organic life form a highly-complex, large-scale interconnected communication network. The genesis and evolution of these macromolecules not only use free energy but also embody the constant storage and transfer of information collected by the living organism through its interaction with its environment. These processes of information storage or, more specifically, preservation of information via DNA replication, and information transfer via message transmission from nucleic acids (DNA and RNA) to proteins form the basis for the continuity of all life on Earth.

The free energy flow process required to sustain organized and complex life forms generating negative (local) entropy is intricately coupled with transferring this negative entropy into information. This increase in information is a result of system replication and information transfer by enzyme and
nucleic acid information-bearing molecules. The information content in a given sequence of proteins depends on the minimum number of specifications needed to characterize a given DNA structure; the more complex the DNA structure, the larger the number of specifications (i.e., microstates) needed to specify or describe the structure. Using Boltzmann’s entropy formula $S = k \log_e W$ numerous physicists have argued that this information entropy increase results in an overall thermodynamic entropy increase.

A general relation between informational entropy and thermodynamic entropy is enigmatic. The two notions are consociated only if they are related explicitly through physical states. Storage of genetic information effects work through the synthesis of DNA polymers, wherein data is stored, encoded, transmitted, and restored, and hence, information content is related to system ordering. However, it is important to note that the semantic component of information (i.e., the meaning of the transmitted message) through binary coding elements is not related to thermodynamic entropy.

Even though the Boltzmann entropy formula gives a measure of disorder or randomness in a closed system, its connection to thermodynamic entropy remains tenuous. Planck was the first to introduce the Boltzmann constant $k$ within the Boltzmann entropy formula to attach physical units in calculating the entropy increase in the nondissipative mixing of two different gasses, thereby relating the Boltzmann (configurational) entropy to the thermodynamic entropy. In this case, the Boltzmann constant relates the average kinetic energy of a gas molecule to its temperature. For more complex systems, however, it is not at all clear how these two entropies are related despite the fact that numerous physicists use the Boltzmann (configurational) and the Clausius (thermodynamic) entropies interchangeably.

The use of the Boltzmann constant by physicists to unify the thermodynamic entropy with the Boltzmann entropy is untenable. Notable exceptions to this pretermission are Shannon [214] and Feynman [215]. In his seminal paper [216], Shannon clearly states that his information entropy, which has the same form as the Boltzmann (configuration) entropy, is relative to the coordinate system that is used, and hence, a change of coordinates will generally result in a change in the value of the entropy. No physical units are ever given to the information entropy and nowhere is the Boltzmann constant mentioned in his work [217]. Similarly, Feynman [215] also never introduces the Boltzmann constant leaving the Boltzmann entropy formula free of any physical units. In addition, he states that the value of the Boltzmann entropy will depend on the artificial number of ways one subdivides the volume of the phase space $W$.

The second law of thermodynamics institutes a progression from complexity (organization and order) to decomplexification (disorder) in the physical universe. While internal system organization and complexity (negative entropy change) through energy dispersal is a necessary condition to polymerize the macromolecules of life, it is certainly not a sufficient condition for the origin of life. The mystery of how thermal, chemical, or solar energy flows from a low-entropy source through simple chemicals leading to the genesis of complex polymer structures remains unanswered.

Natural selection applies to living systems which already have the capacity of replication, and hence, the building of DNA and enzyme architectures by processes other than natural selection is a necessary but not sufficient condition for a prevalent explanation of the origin of life. Living organisms are prodigious due to their archetypical hierarchal complex structures. The coupling mechanism between the flow of free energy, the architectural desideratum in the formation of DNA and protein molecules, and information (configurational) entropy necessary for unveiling the codex of the origin of life remains a mystery.

Even though numerous theoretical and experimental models have been proposed in explaining the coupling mechanism between the flow of free energy, the architectural blueprints for the formation of DNA and protein molecules, and the information necessary for codifying these blueprints, they are all deficient in providing a convincing explanation for the origins of life. These models include mechanisms based on chance [216], neo-Darwinian natural selection [218], inherent self-ordering tendencies in matter [219,220], mineral catalysis [221], and nonequilibrium processes [222,223].
Before unlocking the mystery of the origin of life it is necessary to develop a rigorous unification between thermodynamic entropy, information entropy, and biological organization; these fundamental notions remain disjoint. This is not surprising when one considers how Shannon chose the use of the term entropy for characterizing the amount of information. In particular, he states [224]: “My greatest concern was what to call it. I thought of calling it information, but the word was overly used, so I decided to call it uncertainty. John von Neumann had a better idea, he told me, You should call it entropy, for two reasons. In the first place, your uncertainty function goes by that name in statistical mechanics. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage”.

9. The Second Law, Entropy, Gravity, and Life

The singular and most important notion in understanding the origins of life in the universe is gravity. The inflationary bang that resulted in a gigantic burst of spatial expansion of the early universe followed by gravitational collapse providing the fundamental mechanism for structure formation (e.g., galaxies, stars, planets, accretion disks, black holes, etc.) in the universe was the cradle for the early, low entropy universe necessary for the creation of dissipative structures responsible for the genesis and evolution of life. Inflationary cosmology, wherein initial repulsive gravitational forces and negative pressures provided an outward blast driving space to surge, resulted in a uniform spatial expansion comprised of (almost) uniformly distributed matter [225–227]. In particular, at the end of the inflationary burst, the size of the universe had unfathomably grown and the nonuniformity in the curvature of space had been distended with all density clusters dispersed leading to the nascency of all dissipative structures, including the origins of life, and the dawn of the entropic arrow of time in the universe.

The stars in the heavens provide the free energy sources for heterotrophs through solar photons, whereas the chemical and thermal nonequilibrium states of the planets provide the free energy sources for chemotrophs through inorganic compounds. The three sources of free energy in the universe are gravitational, nuclear, and chemical. Gravitational free energy arises from dissipation in accretion disks leading to angular momentum exchange between heavenly bodies. Nuclear free energy is generated by binding energy per nucleon from strong nuclear forces and forms the necessary gradients for fusion and fission. Chemical free energy is discharged from electrons sliding deeper into electrostatic potential wells and is the energy that heteropathic life avulsed from organic compounds and chemotrophic life draws from inorganic compounds [228].

The origin of these sources of free energy is a direct consequence of the gravitational collapse with solar photons emitted from fusion reactions taking place at extremely high temperatures in the dense centers of stars, and thermal and chemical gradients emanating from volcanic activities on planets. Gravitational collapse is responsible for generating nonequilibrium dissipative structures and structure formation giving rise to temperature, density, pressure, and chemical gradients. This gravitational structure of the universe was spawned from an initial low entropy universe with a nearly homogeneous distribution of matter preceded by inflation and the cosmic horizon problem [212,226,227,229,230].

Given the extreme uniformity of the cosmic microwave background radiation (i.e., the horizon problem) of the universe, it is important to note that the expansion of the universe does not change the entropy of the photons in the universe. This is due to the fact that the expanding universe is adiabatic since the photons within any arbitrary fixed volume of the universe have the same temperature as the surrounding neighborhood of this volume, and hence, heat does not enter or leave the system. This adiabatic expansion implies that the entropy of the photons does not increase. Alternatively, this fact also follows by noting that the entropy of black-body photons is proportional to the number of photons [212], and hence, since the number of photons remains constant for a given fixed volume, so does the entropy.

The question that then remains is, How did inflation effect the low entropy state of the early universe? Since the inflationary burst resulted in a smoothing effect of the universe, one would
surmise that the total entropy decreased, violating the second law of thermodynamics. This assertion, however, is incorrect; even though gravitational entropy decreased, the increase in entropy from the uniform production of matter—in the form of relativistic particles—resulted in a total increase in entropy.

More specifically, gravitational entropy or, equivalently, matter entropy was generated by the gravitational force field. Since matter can convert gravitational potential energy into kinetic energy, matter can coalesce to form clusters that release radiation, and hence, produce entropy. If the matter entropy is low, then the gravitational forces can generate a large amount of entropy via matter clustering. In particular, this matter clustering engenders a nonuniform, nonhomogeneous gravitational field characterized by warps and ripples in spacetime.

The origin of this matter and radiation originated from the release of the enormous potential energy of the Higgs field; the energy field suffused through the universe and connected with the fundamental Higgs boson particle. This field gives rise to mass from fundamental particles in the Higgs field through the Higgs bosons, which contain relative mass in the form of energy—a process known as the Higgs effect.

Thus, false vacuum gauge boson states decayed into a true vacuum through spontaneous symmetry breaking, wherein the release in potential energy from the Higgs effect was (almost) uniformly distributed into the universe cooling and clustering matter. This led to the creation of the cosmic evolution of entropy increase and the gravitational origins of free energy necessary for the genesis and subsistence of life in the universe.

Spontaneous symmetry breaking is a phenomenon where a symmetry in the basic laws of physics appears to be broken. Specifically, even though the dynamical system as a whole changes under a symmetry transformation, the underlying physical laws are invariant under this transformation. In other words, symmetrical system states can transition to an asymmetrical lower energy state but the system Lagrangian is invariant under this action. Spontaneous symmetry breaking of the non-Abelian group SU(2) × U(1), where SU(2) is the special unitary group of degree 2 and U(1) is the unitary group of degree 1, admitting gauge invariance associated with the electroweak force (i.e., the unified description of the electromagnetic and weak nuclear forces) generates the masses of the W and Z bosons, and separates the electromagnetic and weak forces.

However, it should be noted that, in contrast to the U(1) Abelian Lie group involving the gauge electromagnetic field, the isospin transformations under the action of the Lie group SU(2) involving electroweak interactions do not commute, and hence, even though the weak nuclear and electromagnetic interactions are part of a single mathematical framework, they are not fully unified. This is further exacerbated when one includes strong nuclear interactions, wherein the non-Abelian group SU(2) × U(1) is augmented to include the Lie group SU(3) involving charge in quantum chromodynamics to define local symmetry. The continuing quest in unifying the four universal forces—strong nuclear, electromagnetic, weak nuclear, and gravitational—under a single mathematical structure remains irresolute.

The fact that the universe is constantly evolving towards organized clustered matter and gravitational structures is not in violation of the second law. Specifically, even though clustered matter and gravitational structures formed during the gravitational collapse in the universe are more ordered than the early (nearly) uniformly distributed universe immediately following the inflationary burst, the entropy decrease in achieving this ordered structure is offset by an even greater increase in the entropy production generated by the enormous amount of heat and light released in thermonuclear reactions as well as exported angular momenta necessary in the morphing and shaping of these gravitational structures. Organized animate and inanimate hierarchical structures in the universe maintain their highly ordered low entropy states by expelling their internal entropy resulting in a net entropy increase of the universe. Hence, neither gravitational collapse giving rise to order in parts of the universe nor life violate the second law of thermodynamics.
In [9], we established that entropy and time are homeomorphic, and hence, the second law substantiates the arrow of time. Since all living systems constitute nonequilibrium dynamical systems involving dissipative structures, wherein evolution through time is homeomorphic to the direction of entropy increase and the availability of free energy, our existence depends on $dS > 0$. In other words, every reachable and observable state in the course of life, structure, and order in the universe necessitates that $dS > 0$; the state corresponding to $dS = 0$ is an unobservable albeit reachable state. Since a necessary condition for life and order in the universe requires the ability to extract free energy (at low entropy) and produce entropy, a universe that can sustain life requires a state of low entropy. Thus, a maximum entropy universe is neither controllable nor observable; that is, a state of heat death, though reachable, is uncontrollable and unobservable. In a universe where energy is conserved, that is, $dU_{\text{universe}} = 0$, and the cosmic horizon holds, the free energy is given by $dF_{\text{universe}} = -T dS_{\text{universe}}$. Hence, if $dS_{\text{universe}} > 0$, then the flow of free energy continues to create and sustain life. However, as $dS_{\text{universe}} \to 0$ and $T \to 0$, then $dF_{\text{universe}} \to 0$, leading to an eventual demise of all life in that universe.

10. The Second Law, Health, Illness, Aging, and Death

As noted in Section 7, structural variability, complexity, organization, and entropy production are all integral parts of life. Physiological pattern variations, such as heart rate variability, respiratory rate, blood pressure, venous oxygen saturation, and oxygen and glucose metabolism, leading to emergent patterns in our dynamical system physiology are key indicators of youth, health, aging, illness, and eventually death as we transition through life trammelled by the second law of thermodynamics. In physiology, entropy production is closely related to metabolism, which involves the physical and chemical processes necessary for maintaining life in an organism. Since, by the first law of thermodynamics, work energy and chemical energy lead to heat energy, a biological system produces heat, and hence, entropy.

In the central nervous system [231–233], this heat production is a consequence of the release of chemical energy through oxygen metabolism (e.g., oxygen consumption, carbon dioxide, and waste production) and glycolysis (glucose converted to pyruvate) leading to the breakdown of macromolecules to form high-energy compounds (e.g., adenosine diphosphate (ADP) and adenosine triphosphate (ATP)). Specifically, electrochemical energy gradients are used to drive phosphorylation of ADP and ATP through a molecular turbine coupled to a rotating catalytic mechanism. This stored energy is then used to drive anabolic processes through coupled reactions in which ATP is hydrolyzed to ADP and phosphate releasing heat. Additionally, heat production occurs during oxygen intake with metabolism breaking down macromolecules (e.g., carbohydrates, lipids, and proteins) to free high-quality chemical energy to produce work. In both cases, this heat dissipation leads to entropy production [234–236].

Brain imaging studies have found increasing cortical complexity in early life, throughout childhood, and into adulthood [237–239], and then followed by decreasing complexity in the later stages of life [240,241]. Similarly, a rapid rise in entropy production [242] from early life, throughout childhood, and into early adulthood is also seen, followed by a slow decline thereafter [243]. Identical patterns to the rise and fall of entropy production seen through the deteriorating path to aging are also observed in healthy versus ill individuals [236]. Namely, a decrease in cortical complexity (i.e., fractal dimension [244] of the central nervous system) with illness correlating with metabolic decline in glucose metabolism is observed in ill individuals [236,245–247].

This observed reduction in complexity in not limited to the central nervous system and is evidenced in other physiological parameters. For example, the reduction in the degree of irregularity across time scales of the time-series of the heart rate in heart disease patients and the elderly [248–252] as well as the decoupling of biological oscillators during organ failure [253] lead to system decomplexification. Thus, health is correlated with system complexity (i.e., fractal variability) and the ability to effectively and efficiently dissipate energy gradients to maximize entropy production,
whereas illness and aging are associated with decomplexification resulting in decreased entropy production and ultimately death.

The decrease in entropy production and the inability to efficiently dissipate energy gradients results from a decrease in the ability of the living system to extract and consume free energy. This can be due, for example, to diseases such as ischemia, hypoxia, and angina, which severely curtail the body’s ability to efficiently dissipate energy and consequently maximize entropy production, as well as diseases such as renal failure, hypotension, edema, cirrhosis, and diabetes, which result in pathologic retention of entropy.

The loss of system complexity or, equivalently, the loss of complex system dynamics, system connectivity, and fractal variability with illness and age results in the inability of the system to optimally dissipate free energy and maximize entropy production. In other words, vibrant life extracts free energy at low entropy (i.e., high-grade energy) and ejects it at high entropy (i.e., low-grade energy). The overall degree of system complexity provides a measure of the ability of a system to actively adapt and is reflected by the ratio of the maximum system energy output (work expenditure) to the resting energy output (resting work output), in analogy to Carnot’s theorem.

11. The Second Law, Consciousness, and the Entropic Arrow of Time

In this section, we present some qualitative insights using thermodynamic notions that can potentially be useful in developing mechanistic models [254,255] for explaining the underlying mechanisms for consciousness. Specifically, by merging thermodynamics and dynamical systems theory with neuroscience [256,257], one can potentially provide key insights into the theoretical foundation for understanding the network properties of the brain by rigorously addressing large-scale interconnected biological neuronal network models that govern the neuroelectric behavior of biological excitatory and inhibitory neuronal networks.

The fundamental building block of the central nervous system, the neuron, can be divided into three functionally distinct parts, namely, the dendrites, soma (or cell body), and axon. Neurons, like all cells in the human body, maintain an electrochemical potential gradient between the inside of the cell and the surrounding milieu. However, neurons are characterized by excitability—that is, with a stimulus the neuron will discharge its electrochemical gradient and this discharge creates a change in the potential of neurons to which it is connected (the postsynaptic potential). This postsynaptic potential, if excitatory, can serve as the stimulus for the discharge of the connected neuron or, if inhibitory, can hyperpolarize the connected neuron and inhibit firing.

Since the neocortex contains on the order of 100 billion neurons, which have up to 100 trillion connections, this leads to immense system complexity. Computational neuroscience has addressed this problem by reducing a large population of neurons to a distribution function describing their probabilistic evolution, that is, a function that captures the distribution of neuronal states at a given time [258,259].

To elucidate how thermodynamics can be used to explain the conscious-unconscious [260] state transition we focus our attention on the anesthetic cascade problem [261]. In current clinical practice of general anesthesia, potent drugs are administered which profoundly influence levels of consciousness and vital respiratory (ventilation and oxygenation) and cardiovascular (heart rate, blood pressure, and cardiac output) functions. These variation patterns of the physiologic parameters (i.e., ventilation, oxygenation, heart rate variability, blood pressure, and cardiac output) and their alteration with levels of consciousness can provide scale-invariant fractal temporal structures to characterize the degree of consciousness in sedated patients.

The term “degree of consciousness” reflects the intensity of a noxious stimulus. For example, we are often not aware (conscious) of ambient noise but would certainly be aware of an explosion. Thus, the term “degree” reflects awareness over a spectrum of stimuli. For any particular stimulus the transition from consciousness to unconsciousness is a very sharp transition which can be modeled using a very sharp sigmoidal function—practically a step function.
Here, we hypothesize that the degree of consciousness is reflected by the adaptability of the central nervous system and is proportional to the maximum work output under a fully conscious state divided by the work output of a given anesthetized state, once again in analogy to Carnot’s theorem. A reduction in maximum work output (and cerebral oxygen consumption) or elevation in the anesthetized work output (or cerebral oxygen consumption) will thus reduce the degree of consciousness. Hence, the fractal nature (i.e., complexity) of conscious variability is a self-organizing emergent property of the large-scale interconnected biological neuronal network since it enables the central nervous system to maximize entropy production and dissipate energy gradients. Within the context of aging and complexity in acute illnesses, variation of physiologic parameters and their relationship to system complexity and system thermodynamics have been explored in [248–253].

Complex dynamical systems involving self-organizing components forming spatio-temporal evolving structures that exhibit a hierarchy of emergent system properties form the underpinning of the central nervous system. These complex dynamical systems are ubiquitous in nature and are not limited to the central nervous system. As discussed above, such systems include, for example, biological systems, immune systems, ecological systems, quantum particle systems, chemical reaction systems, economic systems, cellular systems, and galaxies, to cite but a few examples. As noted in Section 6, these systems are known as dissipative systems [94,179] and consume free energy and matter while maintaining their stable structure by injecting entropy to the environment.

As in thermodynamics, neuroscience is a theory of large-scale systems wherein graph theory [262] can be used in capturing the (possibly dynamic) connectivity properties of network interconnections, with neurons represented by nodes, synapses represented by edges or arcs, and synaptic efficacy captured by edge weighting giving rise to a weighted adjacency matrix governing the underlying directed dynamic graph network topology [4,5,261]. Dynamic graph topologies involving neuron connections and disconnections over the evolution of time are essential in modeling the plasticity of the central nervous system. Dynamic neural network topologies capturing synaptic separation or creation can be modeled by differential inclusions [263,264] and can provide a mechanism for the objective selective inhibition of feedback connectivity in association with anesthetic-induced unconsciousness.

In particular, recent neuroimaging findings have shown that the loss of top-down (feedback) processing in association with loss of consciousness observed in electroencephalographic (EEG) signals and functional magnetic resonance imaging (fMRI) is associated with functional disconnections between anterior and posterior brain structures [265–267]. These studies show that topological rather than network connection strength of functional networks correlate with states of consciousness. In particular, anesthetics reconfigure the topological structure of functional brain networks by suppressing midbrain-pontine areas involved in regulating arousal leading to a dynamic network topology. Hence, changes in the brain network topology are essential in capturing the mechanism of action for the conscious-unconscious transition rather than network connectivity strength during general anesthesia.

In the central nervous system billions of neurons interact to form self-organizing nonequilibrium structures. However, unlike thermodynamics, wherein energy spontaneously flows from a state of higher temperature to a state of lower temperature, neuron membrane potential variations occur due to ion species exchanges which evolve from regions of higher chemical potentials to regions of lower chemical potentials (i.e., Gibbs’ chemical potential [268]). This evolution does not occur spontaneously but rather requires a hierarchical (i.e., hybrid) continuous-discrete architecture for the opening and closing of specific gates within specific ion channels.

The physical connection between neurons occurs in the synapse, a small gap between the axon, the extension of the cell body of the transmitting neuron, and the dendrite, the extension of receiving neuron. The signal is transmitted by the release of a neurotransmitter molecule from the axon into the synapse. The neurotransmitter molecule diffuses across the synapse, binds to a postsynaptic receptor membrane protein on the dendrite, and alters the electrochemical potential of the receiving neuron.
The time frame for the synthesis of neurotransmitter molecules in the brain far exceeds the time scale of neural activity, and hence, within this finite-time frame neurotransmitter molecules are conserved.

Embedding thermodynamic state notions (i.e., entropy, energy, free energy, chemical potential, etc.) within dynamical neuroscience frameworks [4,5,261] can allow us to directly address the otherwise mathematically complex and computationally prohibitive large-scale brain dynamical models [4,5]. In particular, a thermodynamically consistent neuroscience model would emulate the clinically observed self-organizing, spatio-temporal fractal structures that optimally dissipate free energy and optimize entropy production in thalamocortical circuits of fully conscious healthy individuals. This thermodynamically consistent neuroscience framework can provide the necessary tools involving semistability, synaptic drive equipartitioning (i.e., synchronization across time scales), free energy dispersal, and entropy production for connecting biophysical findings to psychophysical phenomena for general anesthesia.

The synaptic drive accounts for pre- and postsynaptic potentials in a neural network to give a measure of neural activity in the network. In particular, the synaptic drive quantifies the present activity (via the firing rate) along with all previous activity within a neural population appropriately scaled (via a temporal decay) by when a particular firing event occurred. Hence, the synaptic drive provides a measure of neuronal population activity that captures the influence of a given neuron population on the behavior of the network from the infinite past to the current time instant. For details see [4,5].

System synchronization refers to the fact that the dynamical system states achieve temporal coincidence over a finite or infinite time horizon, whereas state equipartitioning refers to the fact that the dynamical system states converge to a common value over a finite or infinite time horizon. Hence, both notions involve state agreement in some sense. However, equipartitioning involves convergence of the state values to a constant state, whereas synchronization involves agreement over time instants. Thus, equipartitioning implies synchronization; however, the converse is not necessarily true. It is only true in so far as consensus is interpreted to hold over time instants.

The EEG signal does not reflect individual firing rates but rather the synchronization of a large number of neurons (organized in cortical columns) generating macroscopic currents. Since there exists EEG activity during consciousness, there must be some degree of synchronization when the patient is awake. However, during an anesthetic transition the predominant frequency changes, and the dynamical model should be able to predict this frequency shift.

In particular, we hypothesize that as the model dynamics transition to an anesthetized state, the system will involve a reduction in system complexity—defined as a reduction in the degree of irregularity across time scales—exhibiting synchronization of neural oscillators (i.e., thermodynamic energy equipartitioning). This would result in a decrease in system energy consumption (myocardial depression, respiratory depression, hypoxia, ischemia, hypotension, vasodilation), and hence, a decrease in the rate of entropy production. In other words, unconsciousness is characterized by system decomplexification, which is manifested in the failure to develop efficient mechanisms to dissipate energy thereby pathologically retaining higher internal (or local) entropy levels.

The human brain is isothermal and isobaric [269], that is, the temperatures of the subnetworks of the brain are equal and remain constant, and the pressure in each subnetwork also remains constant. The human brain network is also constantly supplied with a source of (Gibbs) free energy provided by chemical nourishment of the blood to ensure adequate cerebral blood flow and oxygenation, which involves a blood concentration of oxygen to ensure proper brain function. Information-gathering channels of the blood also serve as a constant source of free energy for the brain. If these sources of free energy are degraded, then internal (local) entropy is produced.

In the transition to an anesthetic state, complex physiologic work cycles (e.g., cardiac respiratory pressure-volume loops, mitochondrial ATP production) necessary for homeostasis follow regressive diffusion and chemical reaction paths that degrade free energy consumption and decrease the rate of entropy production. Hence, in an isolated large-scale network (i.e., a network with no energy
exchange between the internal and external environment) all the energy, though always conserved, will eventually be degraded to the point where it cannot produce any useful work (e.g., oxygenation, ventilation, heart rate stability, organ function).

In this case, all motion (neural activity) would cease leading the brain network to a state of unconsciousness (semistability) wherein all or partial [4,5,270] subnetworks will possess identical energies (energy equipartition or synchronization) and, hence, internal entropy will approach a local maximum or, more precisely, a saddle surface in the state space of the process state variables. Thus, the transition to a state of anesthetic unconsciousness involves an evolution from an initial state of high (external) entropy production (consciousness) to a temporary saddle state characterized by a series of fluctuations corresponding to a state of significantly reduced (external) entropy production (unconsciousness).

In contrast, in a healthy conscious human, entropy production occurs spontaneously and, in accordance with Jaynes’ maximum entropy production principle [205,206], energy dispersal is optimal leading to a maximum entropy production rate. Hence, low entropy in (healthy) human brain networks is synonymous with consciousness and the creation of order (negative entropy change) reflects a rich fractal spatio-temporal variability which, since the brain controls key physiological processes such as ventilation and cardiovascular function, is critical for delivering oxygen and anabolic substances as well as clearing the products of catabolism to maintain healthy organ function.

In accordance with the second law of thermodynamics, the creation and maintenance of consciousness (i.e., internal order—negative entropy change) is balanced by the external production of a greater degree of (positive) entropy. This is consistent with the maxim of the second law of thermodynamics as well as the writings of Kelvin [41], Gibbs [271,272], and Schrödinger [207] in that the creation of a certain degree of life and order in the universe is inevitably coupled by an even greater degree of death and disorder [3,9].

In a network thermodynamic model of the human brain, consciousness can be equated to the brain dynamic states corresponding to a low internal (i.e., local) system entropy. In [3,9], we have shown that the second law of thermodynamics provides a physical foundation for the arrow of time. In particular, we showed that the existence of a global strictly increasing entropy function on every nontrivial network thermodynamic system trajectory establishes the existence of a completely ordered time set that has a topological structure involving a closed set homeomorphic to the real line, which establishes the emergence of the direction of time flow.

Thus, the physical awareness of the passage of time (i.e., chronognosis—κρόνον γνωσις) is a direct consequence of the regressive changes in the continuous rate of entropy production taking place in the brain and eventually leading (in finite time) to a state of no further entropy production (i.e., death). Since these universal regressive changes in the rate of entropy production are spontaneous, continuous, and decrease in time with age, human experience perceives time flow as unidirectional and nonuniform. In particular, since the rate of time flow and the rate of system entropy regression (i.e., free energy consumption) are bijective (i.e., one-to-one and onto), the human perception of time flow is subjective [273].

During the transition to an anesthetized state, the external and internal free energy sources are substantially reduced or completely severed in part of the brain leading the human brain network to a semistable state corresponding to a state of local saddle (stationary) high entropy. Since all motion in the state variables (synaptic drives) ceases in this unconscious (synchronized) state, our index for the passage of time vanishes until the anesthetic wears off allowing for an immediate increase of the flow of free energy back into the brain and other parts of the body. This, in turn, gives rise to a state of consciousness, wherein system entropy production is spontaneously resumed and the patient takes in free energy at low entropy and excretes it at high entropy. Merging system thermodynamics with mathematical neuroscience can provide a mathematical framework for describing conscious-unconscious state transitions as well as developing a deeper understanding on how the arrow of time is built into the very fabric of our conscious brain.
12. Relativistic Mechanics

Classical thermodynamics as well as the dynamical systems framework formulation of thermodynamics presented in [3] are developed for systems that are assumed to be at rest with respect to a local observer and in the absence of strong gravitational fields. To effectively address the universality of thermodynamics and the arrow of time to cosmology, the dynamical systems framework of thermodynamics developed in [3] needs to be extended to thermodynamic systems which are moving relative to a local observer moving with the system and a stationary observer with respect to which the system is in motion. In addition, the gravitational effects on the thermodynamics of moving systems need to also be considered.

Our universe is the quintessential large-scale dynamical system, and classical physics has provided a rigorous underpinning of its governing laws that are consistent with human experience and intuition. Newtonian mechanics and Lagrangian and Hamiltonian dynamics provide a rigorous mathematical framework for classical mechanics—the oldest of the physical sciences. The theory of Newtonian mechanics embodies the science of kinematics and dynamics of moving bodies and embedded within its laws is the Galilean principle of relativity, which states that the laws of mechanics are equally valid for all observers moving uniformly with respect to each other.

This invariance principle of mechanics, which was known from the time of Galileo, postulated that the laws of mechanics are equivalent in all inertial (i.e., nonaccelerating) frames. In other words, all inertial observers undergoing uniform relative motion would observe the same laws of mechanics involving the conservation of energy, conservation of linear momentum, conservation of angular momentum, etc., while possibly measuring different values of velocities, energies, momenta, etc.

In the late seventeenth century, Newton went on to make his greatest discovery—the law of universal gravitation. He postulated that gravity exists in all bodies universally, and showed that bodies attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of their separation. He maintained that this universal principle pervades the entire universe affecting everything in it, and its influence is direct, immediate, and definitive. With this law of universal gravitation, Newton demonstrated that terrestrial physics and celestial physics can be derived from the same basic principles and are part of the same scientific discipline.

Even though the laws of classical (i.e., Newtonian) mechanics provide an isomorphism between an idealized, exact description of the Keplerian motion of heavenly bodies in our solar system deduced from direct astronomical observations, wherein the circular, elliptical, and parabolic orbits of these bodies are no longer fundamental determinants of motion, the integral concepts of space and time necessary in forming the underpinning of classical mechanics are flawed resulting in an approximation of the governing universal laws of mechanics. In particular, in his *Principia Mathematica* [166] Newton asserted that space and time are absolute and unchanging entities in the fabric of the cosmos.

Specifically, he writes: “I do not define time, space, and motion, as being well known to all ... Absolute, true and mathematical time, of itself and from its own nature, flows equably without relation to anything external, and by another name is called duration”. He goes on to state: “Absolute space, in its own nature, without reference to anything external, remains always similar and unmovable”. The absolute structure of time and space of Newtonian mechanics asserts that a body being at rest is at rest with respect to absolute space; and a body is accelerating when it is accelerating with respect to absolute space.

In classical mechanics, Kepler, Galileo, Newton, and Huygens sought to completely identify the fundamental laws that would define the kinematics and dynamics of moving bodies. These laws have become known as Newton’s laws of motion and in the seventeenth through the nineteenth centuries the physical approach of classical mechanics was replaced by mathematical theories involving abstract theoretical structures by giants such as Euler, Lagrange, and Hamilton.

However, what has become known as the most significant scientific event of the nineteenth century is Maxwell’s discovery of the laws of electrodynamics, which sought to do the same thing for electromagnetism as what Newton’s laws of motion did for classical mechanics. Maxwell’s equations...
completely characterize the electric and magnetic fields arising from distributions of electric charges and currents, and describe how these fields change in time [274]. His equations describe how changing magnetic fields produce electric fields; assert the nonexistence of magnetic monopoles; and describe how electric and magnetic fields interact and propagate [275].

Maxwell’s mathematical grounding of electromagnetism included Gauss’ flux theorem [276], relating the distribution of electric charge to the resulting electric field, and cemented Faraday’s profound intuitions based on decades of experimental observations [277]. Faraday was the first to recognize that changes in a magnetic force produces electricity, and the amount of electricity that is produced is proportional to the rate of change of the magnetic force. In Maxwell’s equations the laws of Gauss and Faraday remain unchanged. However, Ampère’s theorem—asserting that magnetic circulation along concentric paths around a straight wire carrying a current is unchanged and is proportional to the current through the permittivity of free space—is modified to account for varying electric fields through the introduction of a displacement current [278]. As in classical mechanics, wherein Euler provided the mathematical formulation of Newton’s $F = ma$ [166,279–283], Maxwell provided the mathematical foundation of Faraday’s researches which lacked mathematical sophistication.

The theory of electromagnetism explained the way a changing magnetic field generates an electric current, and conversely the way an electric field generates a magnetic field. This coupling between electricity and magnetism demonstrated that these physical entities are simply different aspects of a single physical phenomenon— electromagnetism—and not two distinct phenomena as was previously thought. This unification between electricity and magnetism further led to the realization that an electric current through a wire produces a magnetic field around the wire, and conversely a magnetic field induced around a closed-loop is proportional to the electric current and the displacement current it encloses. Hence, electricity in motion results in magnetism, and magnetism in motion produces electricity. This profound insight led to experiments showing that the displacement of electricity is relative; absolute motion cannot be determined by electrical and magnetic experiments.

In particular, in the mid-nineteenth century scientists performed an electrical and magnetic experiment, wherein the current in a conductor moving at a constant speed with respect to the magnet was calculated with respect to the reference frame of the magnet and the reference frame of the conductor. In either case, the current was the same indicating that only relative motion matters and there is no absolute standard of rest. Specifically, the magnet displaced through a loop of wire at a fixed speed relative to the wire produced the same current as when the loop of wire was moved around the magnet in the reverse direction and at the same speed.

In other words, if the relative speed of the magnet and loop of wire is the same, then the current is the same irrespective of whether the magnet or the wire is moving. This implies that there does not exist a state of absolute rest since it is impossible for an observer to deduce whether the magnet or the conductor (i.e., the wire) is moving by observing the current. If the same experiment is repeated in a an inertial frame, then there would be no way to determine the speed of the moving observer from the experiment.

However, it follows from Maxwell’s equations that the charge in the conductor will be subject to a magnetic force in the reference frame of conductor and an electric force in the reference frame of the observer, leading to two different descriptions of the same phenomenon depending on the observer reference frame. This inconsistency is due to the fact that Newtonian mechanics is predicated on a Galilean invariance transformation for the forces driving the charge that produces the current, whereas Maxwellian electrodynamics predicts that the fields generating these forces transform according to a Lorentzian invariance.

Maxwell’s equations additionally predicted that perturbations to electromagnetic fields propagate as electromagnetic waves of electrical and magnetic forces traveling through space at the speed of light. From this result, Maxwell proposed that light is a traveling wave of electromagnetic energy [284,285]. This led to numerous scientists searching for the material medium needed to transmit light waves from
In conforming with the translation invariance principle of Newtonian mechanics requiring the existence of an absolute time and an absolute space, physicists reasoned, by analogy to classical wave propagation theory requiring a material medium for wave transmission, that light travels in a *luminiferous aether*, which defines the state of absolute rest for electromagnetic phenomena.

Numerous experiments [286,287] that followed in attempting to determine a state of absolute rest for electromagnetic phenomena showed that the velocity of light is isotropic. In other words, the speed of light in free space is invariant regardless of the motion of the observer or that of the light’s source. This result highlighted the inconsistency between the principle of relative motion and the principle of the consistency of the speed of light leading to one of science’s most profound repudiation of Occam’s razor.

With the advent of the special theory of relativity [288,289], Einstein dismantled Newton’s notions of absolute time and absolute space. In particular, Einstein developed the special theory of relativity from two postulates. He conjectured that the laws of physics (and not just mechanics) are the same in all inertial (i.e., nonaccelerating) frames; no preferred inertial frame exists, and hence, observers in constant motion are completely equivalent. In other words, all observers moving at uniform speed will arrive at identical laws of physics. In addition, he conjectured that the speed of light [290] \( c \) is the same in all inertial frames, and hence, the speed of light in free space is the same as measured by all observers, independent of their own speed of motion. Einstein’s second postulate asserts that light does not adhere to a particle-motion model, but rather light is an electromagnetic wave described by Maxwell’s field equations, with the caveat that no medium against which the speed of light is fixed is necessary.

The derivation of special relativity makes use of several additional assumptions including spatial homogeneity (i.e., translation invariance of space), spatial isotropy (i.e., rotational invariance of free space), and memorylessness (i.e., past actions cannot be inferred from invariants). These assumptions are assertions about the structure of spacetime. Namely, space and time form a four-dimensional continuum and there exists a *global* inertial frame that covers all space and time, and hence, the geometry of spacetime is flat. (In general relativity, this assumption is replaced with a weaker postulate asserting the existence of a *local* spacetime). Furthermore, Einstein asserted that Newtonian mechanics hold for slow velocities so that the energy-momentum, four-vector in special relativity relating temporal and spatial components collapses to a one energy scalar and one momentum three-vector.

Even though in special relativity velocities, distances across space, and time intervals are relative, *spacetime* is an absolute concept, and hence, relativity theory is an invariance theory. Specifically, space and time are intricately coupled and cannot be defined separately from each other leading to a consistency of the speed of light regardless of the observer’s velocity. Hence, a body moves through space as well as through time, wherein a stationary (i.e., not moving through space) body’s motion is entirely motion through time. As the body speeds away from a stationary observer, its motion through time is transformed into motion through space, slowing down its motion through time for an observer moving with the body relative to the stationary observer. In other words, time slows down for the moving observer relative to the stationary observer leading to what is known as the *time dilation* principle. Thus, special relativity asserts that the combined speed of a body’s motion through spacetime, that is, its motion through space and its motion through time, is always equal to the speed of light. Hence, the maximum speed of the body through space is attained when all the light speed motion through time is transformed into light speed motion through space.

The bifurcation between Newtonian mechanics (absolute space and time) and Einsteinian mechanics (absolute spacetime) is observed when a body’s speed through space is a significant fraction of the speed of light, with time stopping when the body is traveling at light speed through space. Special relativity leads to several counterintuitive observable effects, including time dilation, length contraction, relativistic mass, mass-energy equivalence, speed limit universality, and relativity of simultaneity [291,292].
Time dilation involves the noninvariance between time intervals of two events from one observer to another; length contraction corresponds to motion affecting measurements; and relativity of simultaneity corresponds to nonsynchronization of simultaneous events in inertial frames that are moving relative to one another. Furthermore, in contrast to classical mechanics predicated on the Galilean transformation of velocities, wherein relative velocities add, relativistic velocity addition predicated on the Lorentz transformations results in the speed of light as the universal speed limit, which further implies that no information signal or matter can travel faster than light in a vacuum. This results in the fact that, to all observers, cause will precede effect upholding the principle of causality.

The universal principle of the special theory of relativity is embedded within the postulate that the laws of physics are invariant with respect to the Lorentz transformation characterizing coordinate transformations between two reference frames that are moving at constant velocity relative to each other. In each reference frame, an observer measures distance and time intervals with respect to a local coordinate system and a local clock, and the transformations connect the space and time coordinates of an event as measured by the observer in each frame. At relativistic speeds the Lorentz transformation displaces the Galilean transformation of Newtonian physics based on absolute space and time, and forms a one-parameter group of linear mappings \[293\]. This group is described by the Lorentz transformation wherein a spacetime event is fixed at the origin and can be considered as a hyperbolic rotation of a Minkowski (i.e., flat) spacetime with the spacetime geometry measured by a Minkowski metric.

The special theory of relativity applies to physics with weak gravitational fields wherein the geometric curvature of spacetime due to gravity is negligible. In the presence of a strong gravitational field, special relativity is inadequate in capturing cosmological physics. In this case, the general theory of relativity is used to capture the fundamental laws of physics, wherein a non-Euclidean Riemannian geometry is invoked to address gravitational effects on the curvature of spacetime \[294,295\].

In Riemannian spacetime a four-dimensional, semi-Riemannian manifold \[296\] is used to represent spacetime and a semi-Riemannian metric tensor is used to define the measurements of lengths and angles in spacetime admitting a Levi-Civita connection \[297\], which locally approximates a Minkowski spacetime. In other words, over small inertial coordinates the metric is Minkowskian with vanishing first partial derivatives and Levi-Civita connection coefficients. This specialization leads to special relativity approximating general relativity in the case of weak gravitational fields, that is, for sufficiently small scales and in coordinates of free fall.

Unlike the inverse-square law of Newton’s universal gravitation, wherein gravity is a force whose influence is instantaneous and its source is mass, in the general theory of relativity absolute time and space do not exist, and gravitation is not a force but rather a property of spacetime, with spacetime regulating the motion of matter and matter regulating the curvature of spacetime. Hence, the effect of gravity can only travel through space at the speed of light and is not instantaneous.

The mathematical formulation of the general theory of relativity is characterized by Einstein’s field equations \[294\], which involve a system of ten coupled nonlinear partial differential equations relating the presence of matter and energy to the curvature of spacetime. The solution of Einstein’s field equations provide the components of the semi-Riemannian metric tensor of spacetime describing the geometry of the spacetime and the geodesic trajectories of spacetime.

Einstein’s field equations lead to the equivalence between gravity and acceleration. In particular, the force an observer experiences due to gravity and the force an observer experiences due to acceleration are the same, and hence, if an observer experiences gravity’s influence, then the observer is accelerating. This equivalence principle leads to a definition of a stationary observer as one who experiences no forces. In Newtonian mechanics as well as in special relativity, space and spacetime provide a reference for defining accelerated motion. Since space and spacetime are absolute in Newtonian mechanics and the special theory of relativity, acceleration is also absolute. In contrast, in general relativity space and time are coupled to mass and energy, and hence, they are dynamic and
not unchangeable. This dynamic warping and curving of spacetime is a function of the gravitational field leading to a relational notion of acceleration. In other words, acceleration is relative to the gravitational field [298].

As in the case of special relativity, general relativity also predicts several subtle cosmological effects. These include gravitational time dilation and frequency shifting, light deflection and gravitational time delay, gravitational waves, precession of apsides, orbital decay, and geodetic precession and frame dragging [295, 299]. Gravitational time dilation refers to the fact that dynamic processes closer to massive bodies evolve slower as compared to processes in free space, whereas frequency shifting refers to the fact that light frequencies are increased (blueshifting) when light travels towards an increased gravitational potential (moving towards a gravity well) and decreased (redshifting) when light travels away from a decreased gravitational potential (moving away from a gravity well).

Light deflection refers to the bending of light trajectories in a gravitational field towards a massive body, whereas gravitational time delay refers to the fact that light signals take longer to move through a gravitational field due to the influence of gravity on the geometry of space. General relativity additionally predicts the existence of gravitational waves; ripples in the metric of the spacetime continuum propagating at the speed of light. Finally, general relativity also predicts precession of orbital apsides, orbital decay due to gravitational wave emission, and frame dragging or gravitomagnetic effects (i.e., relativity of direction) in a neighborhood of a rotating mass.

13. Relativistic Thermodynamics

The earliest work on relativistic thermodynamics can be traced back to Einstein [300] and Planck [301], wherein they deduced that entropy is a relativistic invariant and the Lorentz transformation law gives the relativistic temperature change in a moving body as [302, 303].

\[
T' = T \sqrt{1 - \frac{v^2}{c^2}},
\]

(11)

with the heat flux transforming as

\[
dQ' = dQ \sqrt{1 - \frac{v^2}{c^2}},
\]

(12)

where \(T\) and \(dQ\) are the Kelvin temperature and heat flux as measured in the rest system of coordinates (i.e., proper coordinates), \(T'\) and \(dQ'\) are the corresponding temperature and heat flux detected in the moving inertial system, and \(v\) is the inertial frame velocity. Thus, a moving body in the Einstein-Planck theory would appear colder to a moving observer and the heat flux would correspondingly diminish. However, ever since Einstein and Planck attempted to unify special relativity theory with thermodynamics, the Einstein-Planck relativistic thermodynamic theory has been subjected to a heated controversy and considerable debate over the past one hundred years; see for example [304–333].

This is primarily due to fact that thermodynamic variables such as entropy, temperature, heat, and work are global quantities, and hence, they need to be defined with respect to a specific class of hyperplanes in the Minkowski spacetime. Moreover, unlike the thermodynamics of stationary systems, wherein the system energy can be uniquely decomposed into the system internal energy, the energy due to heat (energy) transfer, and the energy due to mechanical work, in relativistic thermodynamics this decomposition is not covariant since heat exchange is accompanied by momentum flow, and hence, there exist nonunique ways in defining heat and work leading to an ambiguity in the Lorentz transformation of thermal energy and temperature.

If the first and second laws of thermodynamics are required to satisfy the principle of relativity, then it is necessary to separate heat transfer from mechanical work for the thermodynamics of moving systems. In other words, a specific definition for heat transfer needs to be decided upon and which further reduces to the classical definition of heat transfer for the thermodynamics of stationary systems. However, in inertially moving thermodynamic systems, the system energy transfer cannot be uniquely separated...
into heat and mechanical work; this arbitrary division in defining heat and mechanical work for relativistic systems has lead to different formulations of relativistic thermodynamics.

The source for the nonuniqueness of how temperature transforms under the Lorentz transformation law stems from the fact that energy and momentum are interdependent in relativity theory. Furthermore, and quite surprisingly, there does not exist a mathematically precise definition of temperature in the literature. Thus, attempting to define temperature (defined only for equilibrium processes in classical thermodynamics) from the spatio-temporal (dynamic) energy-momentum tensor results in a nonunique decomposition for heat and work in relativistic systems. This observation went unnoticed by Einstein and Planck as well as many of the early architects of relativity theory including Tolman [316], Pauli [334], and von Laue [315].

In the later stages of his life, however, Einstein expressed doubt about the thermodynamic transformation laws leading to (11). In a series of letters to Max von Loue in the early 1950s, Einstein changed his view from moving bodies becoming cooler to moving bodies becoming hotter by the Lorentz factor, to finally concluding that “the temperature should be treated in every case as an invariant” [335,336].

A little over five decades after Einstein and Planck formulated their relativistic thermodynamic theory, Ott [306] challenged the Einstein-Planck theory by alleging that the momentum due to heat transfer had been neglected in their formulation, and hence, since conservation of momentum is an integral part of conservation of energy in relativity theory, Ott claimed that their temperature transformation for relativistic thermodynamics was incorrect. Specifically, his formulation postulated the invariance of entropy and asserted that the relativistic temperature transforms as

\[ T' = \frac{T}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \tag{13} \]

with the heat flux transforming as

\[ dQ' = \frac{dQ}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}. \tag{14} \]

In the Ott formulation, a moving body would appear hotter to a moving observer and the heat flux would correspondingly be larger. This formulation was actually arrived at fifteen years earlier by Blanusa [327] but had gone unnoticed in the mainstream physics literature. A few years after Ott’s result, similar conclusions were also independently arrived at by Arzelies [307]. Ever since the Blanusa-Ott formulation of relativistic thermodynamics appeared in the literature, a numerous number of papers followed (see, for example, [304,305,308–314,317–325] and the references therein) adopting the Blanusa-Ott theory as well as asserting new temperature transformation laws.

In an attempt to reconcile the two competing relativistic thermodynamic theories, Landsberg [308,309] maintained that temperature has a real physical meaning only in a laboratory frame (i.e., the frame of reference in which the laboratory is at rest) and asserted that temperature is invariant with respect to the speed of an inertial frame. Furthermore, he asserted that it is impossible to test for thermal equilibrium between two relatively moving inertial systems. Thus, to circumvent the apparent contradiction in the temperature transformation between the Einstein-Planck and Blanusa-Ott theories, Landsberg postulated that

\[ T' = T, \tag{15} \]

with the heat flux transforming as in (12).

Van Kampen [304,305] went one step further by arguing that it is impossible to establish a good thermal contact of a moving body at relativistic speeds, and hence, it is impossible to measure its temperature. This is a blatantly vacuous statement and in essence declares that thermodynamics does not exist in relativity theory. It is just as difficult to measure distance in a fast moving body with
respect to a stationary reference frame. This, however, does not preclude the use of radar methods and the simultaneity concept to compare electromagnetic waves and consequently define length. Alternatively, an optical pyrometer can be used to measure the temperature of a moving body from a distance without the need for establishing thermal contact. In the Van Kampen formulation of relativistic thermodynamics, a covariant thermal energy-momentum transfer four-vector is introduced that defines an inverse temperature four-vector resulting in Lorentz invariance of entropy, heat, and temperature; that is, \( S' = S \), \( dQ' = dQ \), and \( T' = T \). In this case, a moving body would neither appear hotter or colder to a moving observer.

In addition to the aforementioned relativistic temperature transformations, Landsberg and Johns [318] characterize a plethora of possible temperature transformations all of which appear plausible in unifying thermodynamics with special relativity. They conclude that satisfactory temperature measurements in moving thermodynamic systems must lead to \( T' = T \). However, their result leads to a violation of the principle of relativity as applied to the relativistic second law of thermodynamics. Balescu [326] gives a summary of the controversy between relativity and temperature, with a more recent discussion of the relativistic temperature transformation and its ramifications to statistical thermodynamics given in [328–330].

With the notable exception of [323,333], all of the aforementioned relativistic thermodynamic formalisms are based on a Lorentz invariance of the system entropy, that is, \( S' = S \). This erroneous deduction is based on the Boltzmann interpretation of entropy and can be traced back to the original work of Planck [301]. Namely, Planck argued that since entropy corresponds to the number of possible thermodynamic system states, it must be “naturally” Lorentz invariant.

The definition of entropy involving the natural logarithm of a discrete number of system states, however, is confined to thermostatics and does not account for nonequilibrium thermodynamic systems. Furthermore, all of the proposed theories are confined to a unification of thermodynamics with special relativity, and with Tolman [316] providing the most elaborate extension of thermodynamics to general relativity. Moreover, all of the relativistic thermodynamic formalisms are predicated on equilibrium systems characterized by time-invariant tensor densities in different inertial (for special relativity) and noninertial (for general relativity) reference frames.

Because of the relativistic equivalence of mass and energy, the mass associated with heat transfer in the Einstein-Planck relativistic thermodynamic theory is perceived as momentum, and hence, mechanical work by an observer in a moving frame. Alternatively, in the Ott-Arzelies relativistic thermodynamic theory this same work is viewed as part of the heat transfer. Since both theories postulate a Lorentz invariance for entropy, it makes no physical difference which relativistic thermostatic formalism is adopted. In other words, there does not exist an experiment that can uniquely define heat and work in relativistic thermostatic systems.

However, as noted in Section 12, to effectively address the universality of thermodynamics to cosmogony and cosmology involving the genesis and dynamic structure of the cosmos, a relativistic thermodynamic formalism needs to account for nonequilibrium thermodynamic systems. Since, as shown in [3,9], in a dynamical systems formulation of thermodynamics the existence of a global strictly increasing entropy function on every nontrivial trajectory establishes the existence of a completely ordered time set that has a topological structure involving a closed set homeomorphic to the real line, entropy invariance is untenable.

In other words, not only is time and space intricately coupled, but given the topological isomorphism between entropy and time, and Einstein’s time dilation assertion that increasing a body’s speed through space results in decreasing the body’s speed through time, relativistic thermodynamics leads to an entropy dilation principle, wherein the rate of change in entropy increase of a moving system would decrease as the system’s speed increases through space.

More succinctly, motion affects the rate of entropy increase. In particular, an inertial observer moving with a system at a constant velocity relative to a stationary observer with respect to whom the thermodynamic system is in motion will experience a slower change in entropy.
increase when compared to the entropy of the system as deduced by the experimental findings obtained by the stationary observer. This deduction removes some of the ambiguity of relativistic thermodynamic systems predicated on thermostatics, asserts that the Boltzmann constant is not invariant with the speed of the inertial frame, and leads to several new and profound ramifications of relativistic thermodynamics.

### 14. Special Relativity and Thermodynamics

In this section, we use special relativity theory along with the dynamical systems theory of thermodynamics of [9] to develop a thermodynamic theory of moving systems. In particular, we present the relativistic transformations for the thermodynamic variables of volume, pressure, energy, work, heat, temperature, and entropy of thermodynamic systems which are in motion relative to an inertial reference frame. More specifically, we relate the thermodynamic variables of interest in a given inertial frame $\mathcal{F}'$ with the thermodynamic variables as measured in a proper frame $\mathcal{F}$ by a local observer moving with the thermodynamic system in question.

Here we use the thermodynamic relation notions as established in [9] (Chapter 5). Furthermore, for simplicity of exposition, we consider a single subsystem (i.e., compartment) in our analysis; the multicompartment case as delineated in [9] can be treated in an identical manner, with the analogous scalar-vector definitions for compartmental volume, pressure, energy, work, heat, temperature, and entropy.

To develop the relations connecting spatial and temporal measurements between two observers in relative motion, consider the two systems of spacetime coordinates $\mathcal{F}$ and $\mathcal{F}'$ shown in Figure 1. We assume that the inertial frame $\mathcal{F}'$ is moving at a constant velocity $v$ with respect to the proper inertial frame $\mathcal{F}$. For simplicity of exposition, we choose three linearly independent and mutually perpendicular axis, and consider relative motion between the origins of $\mathcal{F}$ and $\mathcal{F}'$ to be along the common $x-x'$ axis so that $v = [v \ 0 \ 0]^T$. This orientational simplification along with the chosen relative velocity of the frames does not affect the physical principles of the results we will derive. The two observers use location measurement devices, which have been compared and calibrated against one another, and time measurement devices (clocks), which have been synchronized and calibrated against one another [291,337]. Furthermore, at $t = 0$ and $t' = 0$, the two origins $O$ and $O'$ of the two reference frames $\mathcal{F}$ and $\mathcal{F}'$ coincide.

Using the principle of relativity, the first and second laws of thermodynamics retain their usual forms in every inertial frame. Specifically, the first law gives

$$\Delta U = -L + Q,$$

which states that the change in energy $\Delta U$ of a thermodynamic system is a function of the system state and involves an energy balance between the heat received $Q$ by the system and the work $L$ done by the system. Furthermore, the second law leads to

$$dS \geq \frac{dQ}{T},$$

which states that the system entropy is a function of the system states and its change $dS$ is bounded from below by the infinitesimal amount of the net heat absorbed $dQ$ or dissipated $-dQ$ by the system at a specific temperature $T$. Note that both the first and second laws of thermodynamics involve statements of changes in energy and entropy, and do not provide a unique zero point for the system energy and entropy.
The primed reference frame \( \mathcal{F}' \) is moving with a constant velocity \( v \) relative to the unprimed reference frame \( \mathcal{F} \) as measured by a stationary observer in \( \mathcal{F} \). Alternatively, by the principle of relativity, a stationary observer in \( \mathcal{F}' \) measures a velocity \(-v\) relative to \( \mathcal{F} \).

The mass-energy equivalence from the theory of relativity, however, gives an additional relationship connecting the change in energy and mass via the relativistic relation

\[
\Delta U = \Delta mc^2, \tag{18}
\]

where \( \Delta m \) denotes the increase in system mass over the system rest mass. Thus, even though the first law of thermodynamics for stationary systems gives only information of changes in system energy content without providing a unique zero point of the energy content, the relativistic energy relation

\[
U = mc^2 \tag{19}
\]

infers that a point of zero energy content corresponds to the absence of all system mass. As discussed in [9], the zero point of entropy content is provided by the third law of thermodynamics, which associates a zero entropy value to all pure crystalline substances at an absolute zero temperature.

To develop the relativistic Lorentz transformations for the thermodynamic variables, we assume that the thermodynamic system is at rest relative to the proper inertial reference frame \( \mathcal{F} \). Furthermore, the thermodynamic quantities of internal energy \( U' \), heat \( Q' \), work \( W' \), volume \( V' \), pressure \( p' \), temperature \( T' \), and entropy \( S' \) are measured or deduced with respect to the inertial reference frame \( \mathcal{F}' \) moving at constant velocity \( u \), with magnitude \( u \), relative to the reference frame \( \mathcal{F} \). The Lorentz transformation equations for the thermodynamic quantities in \( \mathcal{F}' \) moving with uniform velocity relative to \( \mathcal{F} \), as measured by a stationary observer in \( \mathcal{F} \), should require that the laws of thermodynamics are valid relativistically, and hence, are invariant under the Lorentz transformation equations. Thus, the first and second laws of thermodynamics in the \( \mathcal{F}' \) frame should be equivalent in form to a proper coordinate frame \( \mathcal{F} \), wherein the thermodynamic system is at rest.

Here, we develop the transformation equations in a form relating the thermodynamic quantities in the reference frame \( \mathcal{F}' \) with their respective quantities as measured or deduced in the proper frame \( \mathcal{F} \) by a local observer moving with the thermodynamic system. In other words, we ask the question: If a body is moving with a constant velocity \( u \) relative to a rest (i.e., proper) system of coordinates \( \mathcal{F} \), then how are its thermodynamic quantities detected in \( \mathcal{F}' \) related to the thermodynamic quantities in \( \mathcal{F} \) as measured or deduced in the rest system of coordinates \( \mathcal{F} \)?
First, we relate the compartmental volume $V'$ of a thermodynamic system moving with uniform velocity $u$, as measured by a stationary observer in $\mathcal{F}$, to the compartment volume $V$ as measured by the observer in $\mathcal{F}$. Here we assume that the thermodynamic system state is defined by energy and volume or temperature and pressure with the material substance in the compartment exerting equal pressure in all directions and not supporting shear. Since there is no disagreement in the length measurement between the two system of coordinates at right angles to the line of motion and, by Lorentz contraction, a body’s length will measure shorter in the direction of motion, it follows from the length contraction principal that

$$V' = V \sqrt{1 - \left(\frac{u}{c}\right)^2},$$

where $V$ is the compartmental volume as measured in the proper coordinate frame $\mathcal{F}$.

Next, to relate the pressure $p'$ in the compartment of the moving thermodynamic system relative to the compartmental pressure $p$ as measured in the proper coordinates, recall that pressure is defined as the applied force per unit area. Now, since the relative motion between the origins of $\mathcal{F}$ and $\mathcal{F}'$ is along the common $x$-$x'$ axis and the forces $F_x, F_y,$ and $F_z$ acting on the surfaces of the compartmental system are perpendicular to the $(x, y, z)$ and $(x', y', z')$ axes, it follows from the relativistic force transformation equations [9,291] that

$$F'_x = F_x,$$

$$F'_y = F_y \sqrt{1 - \left(\frac{u}{c}\right)^2},$$

$$F'_z = F_z \sqrt{1 - \left(\frac{u}{c}\right)^2},$$

where $F_x, F_y,$ and $F_z$ are the forces acting on the surfaces of the compartmental system as measured in the reference frame $\mathcal{F}$. Now, since the compartment area in the direction of motion will not be affected by length contraction, whereas the areas perpendicular to the $y'$ and $z'$ axes will undergo a length contraction of a ratio of $\sqrt{1 - \left(\frac{u}{c}\right)^2}$ to 1, it follows form (21)–(23) and the definition of pressure as force per unit area that

$$p' = p,$$

where $p$ is the compartmental pressure as measured in the proper coordinate frame $\mathcal{F}$.

Next, we use the momentum density in spacetime describing the density of energy flow [9] to obtain an expression for the energy of the moving thermodynamic compartment. We assume that the mass density and momentum density of the compartment as measured by an inertial observer moving with the compartment are $\rho'$ and $p'_{d'}$, respectively. Here, we start the system at a state of rest and calculate the work necessary to drive the system to a given velocity $u$. First, we assume that the acceleration to the given velocity $u$ takes place adiabatically, wherein changes in the velocity take place without the flow of heat or any changes in the internal system conditions as measured by a local observer. Then, in analogy to Planck’s formulation [301], we use the relation between the mass, energy, and momentum to develop an expression for the momentum of the moving compartment. Specifically, because of the mass-energy equivalence of special relativity, the system momentum will in general change as function of energy even if the system velocity remains constant.

In particular, the total momentum density of the compartment is given by the sum of the momentum density of the moving mass in the compartment and the momentum density associated with the energy flow (i.e., power) density resulting from the work done on the moving material by the stress forces (i.e., pressure) acting on the moving compartmental walls giving

$$p'_d = \rho' u + \frac{p'}{c^2} u.$$


Now, using $\rho' = \frac{\rho'}{\rho} = \frac{\rho' u'}{c^2}$, it follows from (25) that

$$p'_d = \left( \frac{U' + p' V'}{c^2} \right) u,$$

and hence, the momentum of the moving compartment with volume $V'$ is given by

$$p' = \left( \frac{U' + p' V'}{c^2} \right) u.$$  (27)

Now, the force $F'$ generated by accelerating the compartment from a state of rest (i.e., zero velocity) to the final velocity $u$ is given by the rate of change of momentum, and hence,

$$F' = \frac{d}{dt} p' = \frac{d}{dt} \left[ \left( \frac{U' + p' V'}{c^2} \right) u \right].$$  (28)

Next, to compute the work done along with the energy increase when the compartment moves from a state of rest to a velocity $u$, note that the rate of change in energy is given by

$$\frac{dU'}{dt'} = F'T' - p'dV'/dt',$$

where $F'T'$ denotes the power generated by the applied force $F'$ producing the acceleration and $-p'dV'/dt'$ denotes the power dissipated by stress forces (i.e., pressure) acting on the compartment volume, which, due to the Lorentz contraction, is decreasing in its length in the direction of motion, that is, in the $x$-$x'$ axis.

Next, using (20) and (24), noting that $p = p'$ and $V$ are constants, and substituting (28) for $F'$ in (29) yields

$$\frac{dU'}{dt'} = \frac{d}{dt'} \left[ \left( \frac{U' + p' V'}{c^2} \right) u \right] - p'dV'/dt'$$

or, equivalently,

$$\left( 1 - \frac{u^2}{c^2} \right) \frac{d}{dt'} \left( U' + p' V' \right) = \frac{U' + p' V'}{c^2} \frac{du}{dt'}.$$  (31)

Now, rearranging (31) and integrating from a zero initial velocity at time $t = t' = 0$ to a final terminal velocity $u$ at time $t'$ yields

$$\int_0^{t'} \frac{d}{dt'} \left( U' + p' V' \right) = \int_0^u \frac{\sigma}{c^2 - \sigma^2} d\sigma,$$  (32)

which implies

$$\log_e \left( \frac{U' + p' V'}{U + p V} \right) = \frac{1}{2} \log_e \frac{c^2 - \frac{1}{2} \left( c^2 - u^2 \right)}{c^2 - \frac{1}{2} \left( c^2 - u^2 \right)},$$  (33)

or, equivalently,

$$\log_e \left( \frac{U' + p' V'}{U + p V} \right) = \log_e \left( \frac{c}{\sqrt{c^2 - u^2}} \right).$$  (34)
Hence,

\[ U' + p'V' = \frac{U + pV}{\sqrt{1 - (u/c)^2}}. \]  \hfill (35)

Finally, using (20) and (24), (35) yields

\[ U' = \frac{U + pV (u^2/c^2)}{\sqrt{1 - (u/c)^2}}. \]  \hfill (36)

We now turn our attention to the differential work \( dW' \) done by the compartmental system through the external applied force necessary to maintain a constant velocity \( u \). Here, recall that it follows from relativistic dynamics that the equivalence between mass and energy couples the conservation of energy with the conservation of momentum. Hence, the momentum of the compartmental system can change even if the velocity of the system is constant when the compartmental energy changes.

Thus, to maintain a constant velocity, it is necessary to include an external force that balances the differential work done by the stress forces due to pressure. Specifically, the differential work done by the compartmental system at constant velocity is given by

\[ dW' = p'dV' - u^Tdp', \]  \hfill (37)

where \( p'dV' \) denotes the differential work done by the compartmental system due to stress forces (i.e., pressure) and \( u^Tdp' \) denotes the differential work done to the compartmental system by an external force necessary to maintain the constant velocity \( u \).

Next, using (27) it follows from (37) that

\[
\begin{align*}
dW' &= p'dV' - u^Tdp' \\
&= \frac{U' + p'V'}{c^2} u \\
&= p'dV' - \frac{u^2/c^2}{c^2} (U' + p'V'). \\
&= p'dV' - \frac{u^2/c^2}{c^2} (U' + p'V').
\end{align*}
\]  \hfill (38)

Now, using (20), (24), and (36) yields

\[
\begin{align*}
dW' &= p'dV' - \frac{u^2/c^2}{c^2} (U' + p'V') \\
&= dW' \frac{\sqrt{1 - (u/c)^2}}{\sqrt{1 - (u/c)^2}} - \frac{u^2/c^2}{c^2} (U + pV),
\end{align*}
\]  \hfill (39)

or, equivalently,

\[
\begin{align*}
dW' &= dW' \frac{\sqrt{1 - (u/c)^2}}{\sqrt{1 - (u/c)^2}} - \frac{u^2/c^2}{c^2} (U + pV),
\end{align*}
\]  \hfill (40)

which gives an expression for the differential work done by the moving compartment in terms of the differential work, internal energy, pressure, and volume as measured in the proper coordinate frame \( F \).

Using the principle of relativity, the first law of thermodynamics holds for all inertial frames, and hence, it follows from (16) that

\[ dQ' = dU' + dW' \]  \hfill (41)

and

\[ dQ = dU + dW. \]  \hfill (42)
Next, using (36) it follows that
\[ dU' = \frac{dU + d(pV) \left( \frac{u^2}{c^2} \right)}{\sqrt{1 - \left( \frac{u}{c} \right)^2}}, \] (43)
and hence, (43), (39), and (41) imply
\[ dQ' = \frac{dU + d(pV) \left( \frac{u^2}{c^2} \right)}{\sqrt{1 - \left( \frac{u}{c} \right)^2}} + dW \sqrt{1 - \left( \frac{u}{c} \right)^2} \]
\[-\frac{u^2/c^2}{\sqrt{1 - \left( \frac{u}{c} \right)^2}} \left[ dU + d(pV) \right] \]
\[ = (dU + dW) \sqrt{1 - \left( \frac{u}{c} \right)^2}. \] (44)

Now, using (42) it follows that
\[ dQ' = dQ \sqrt{1 - \left( \frac{u}{c} \right)^2}. \] (45)

Finally, we turn our attention to relativistic entropy and temperature. Given the topological isomorphism between entropy and time, and Einstein’s time dilation principle asserting that a body’s measured time rate will measure faster when it is at rest relative to the time rate of a moving body, it follows, in analogy to the relativistic time dilation equation
\[ dt = \frac{dt'}{\sqrt{1 - \left( \frac{v}{c} \right)^2}}, \] (46)
and assuming a linear isomorphism between entropy and time, that
\[ dS' = dS \sqrt{1 - \left( \frac{u}{c} \right)^2}. \] (47)

Now, since by the principle of relativity the second law of thermodynamics must hold for all inertial frames, it follows from (17) that
\[ dS' \geq \frac{dQ'}{T'}, \] (48)
and
\[ dS \geq \frac{dQ}{T}. \] (49)
Thus, using (45) and (47), and (48) and (49), it immediately follows that
\[ T' = T, \] (50)
which shows that temperature is invariant with the speed of the inertial reference frame.

Our formulation here presents a key departure from virtually all of the existing relativistic thermodynamic formulations in the literature which assume that entropy is Lorentz invariant [323,333,338]. Furthermore, it is important to note that even though we assumed that the relational change between the proper and nonproper entropy transforms exactly as the relational change between the proper and nonproper time interval, entropy dilation and time dilation with respect to a stationary observer differ by an invariant factor. In other words, motion affects the rate at which entropy increases in the same way as motion affects the rate at which time increases. Specifically, the change in system entropy slows down with motion analogously to moving clocks running slow.

Thus, since time dilation with respect to a stationary observer in relativity theory applies to every natural phenomena involving the flow of time, biological functions such as heart rate, pulse rate, cell apoptosis, etc. would run at a slower rate with motion resulting in a slowdown of aging. If motion
affects the rate of a biological clock in the same way it affects a physical clock, then one would expect motion to also affect entropy, which is responsible for the breakdown of every living organism and nonliving organized system in nature, in the same way. In other words, motion affects the rates of both physical and biological clocks leading to an entropy dilation with respect to a stationary observer relative to a moving observer.

This observation is completely consistent with the correct interpretation of the twin paradox thought experiment [291]. Recall that the perceived logical contradiction in Einstein’s twin paradox thought experiment follows from the notion that in relativity theory either twin can be regarded as the traveler, and hence, depending on which twin is viewed as the stationary observer, time dilation leads to each twin finding the other younger. This seemingly paradoxical conclusion is predicated on the erroneous assumption that both reference frames are symmetrical and interchangeable. However, the fact that one of the twins would have to execute a round-trip trajectory implies that the reference frame attached to the traveling twin is not inertial since the traveling twin will have to accelerate and decelerate over their journey. This leads to a nonsymmetrical situation wherein one twin is always in an inertial frame whereas the other twin is not.

One can also argue that since general relativity stipulates that all observers (inertial and noninertial) are equivalent, then the twin paradox is not resolved. However, because of the principle of equivalence one can easily show that this reformulation of the twin paradox is also unsound. Specifically, if we assume that the Earth moves away from the traveling twin, then the entire universe must move away with the Earth; and in executing a round-trip trajectory, the whole universe would have to accelerate, decelerate, and then accelerate again. This accelerating universe results in a gravitational field in the reference frame attached to the nontraveling twin slowing down time, whereas the traveling twin making the round trip journey feels no accelerations. When computing the frequency shifts of light in this gravitational field, we arrive at the same conclusion as we would have obtained if we were to assume that the traveling twin executed the round trip journey.

In arriving at (50) we assumed that the homeomorphic relationship between the entropy dilation and time dilation is linear, and hence, using (46) we deduced (47). The fact, however, that there exists a topological isomorphism between entropy and time does not necessarily imply that the relativistic transformation of the change in entropy scales as the relativistic transformation for time dilation. In general, (47) can have the more general form

\[
\frac{dS'}{dS} = f(\beta), \quad (51)
\]

where \( f : [0, 1] \to \mathbb{R}_+ \) is a continuous function of \( \beta \overset{\Delta}{=} v/c \). In this case, since (45) holds, the temperature will have to scale as

\[
\frac{T'}{T} = g(\beta), \quad (52)
\]

where \( g : [0, 1] \to \mathbb{R}_+ \) is continuous.

Thus, it follows from (45), (48), (49), (51), and (52) that

\[
\frac{dS'}{dS} = f(\beta) \geq \frac{dQ'}{T'} = \frac{dQ}{g(\beta)T}, \quad (53)
\]

and hence,

\[
dS \geq \frac{dQ \sqrt{1 - \beta^2}}{T f(\beta)g(\beta)}. \quad (54)
\]

Now, since (49) holds, it follows from (54) that

\[
f(\beta)g(\beta) = \sqrt{1 - \beta^2}. \quad (55)
\]
The indeterminacy manifested through the extra degree-of-freedom in the inability to uniquely specify the functions $f(\beta)$ and $g(\beta)$ in (55) is discussed in Section 15.

From the above formulation, the thermodynamics of moving systems is characterized by the thermodynamic equations involving the thermodynamic transformation variables given by

$$V' = V \sqrt{1 - (u/c)^2},$$  \hspace{2cm} (56)  
$$p' = p,$$  \hspace{2cm} (57)  
$$U' = \frac{U + pV (u^2/c^2)}{\sqrt{1 - (u/c)^2}},$$  \hspace{2cm} (58)  
$$dW' = dW \sqrt{1 - (u/c)^2} - \frac{u^2/c^2}{\sqrt{1 - (u/c)^2}} d(U + pV),$$  \hspace{2cm} (59)  
$$dQ' = dQ \sqrt{1 - (u/c)^2},$$  \hspace{2cm} (60)  
$$dS' = f(\beta) dS,$$  \hspace{2cm} (61)  
$$T' = g(\beta) T.$$  \hspace{2cm} (62)

Since these equations constitute a departure from the thermodynamic equations for stationary systems and involve second-order or higher factors of $u/c$, it is clear that it is impossible to experimentally verify relativistic thermodynamic theory given the present state of our scientific and technological capabilities.

15. Relativity, Temperature Invariance, and the Entropy Dilation Principle

In Section 14, we developed the relativistic transformations for the thermodynamic variables of moving systems. These equations were given by (56)–(62) and provide the relationship between the thermodynamic quantities as measured in the rest system of coordinates (i.e., proper coordinates) to the corresponding quantities detected or deduced in an inertially moving system. However, as discussed in Section 13, there exist several antinomies between the relativistic transformations of temperature in the literature that have not been satisfactorily addressed and that have profound ramifications in relativistic thermodynamics and cosmology. Furthermore, with the notable exception of our formulation along with the suppositions in [323,333], all of the relativistic transformations of temperature in the literature are predicated on the erroneous conjecture that entropy is a Lorentz invariant. The topological isomorphism between entropy and time shows that this conjecture is invalid.

Even though the authors in [323,333] arrive at similar conclusions involving a Lorentz invariance of temperature and a relativistic transformation of entropy, their hypothesis is not based on the mathematical formalism established in [3,9] yielding a topological equivalence between entropy and time, and leading to an emergence between the change in entropy and the direction of time flow. Specifically, the author in [323] conjectures without any proof (mathematical or otherwise) that temperature is invariant with speed.

The main argument provided alleges that if the universe is expanding with distant galaxies moving away from each other at relatively high speeds and (11) holds, then these galaxies should be cold and invisible. Alternatively, if (13) holds, then these galaxies should be infinitely hot and bright. Since neither of these are observable phenomena in the universe, [323] claims that temperature must be invariant with speed.

This assertion is arrived at through a series of physical arguments based on phenomenological considerations that seem to lead to the invariance of temperature under a Lorentz transformation. The main argument provided alleges that if the universe is expanding with distant galaxies moving away from each other at relatively high speeds and (11) holds, then these galaxies should be cold and invisible. Alternatively, if (13) holds, then these galaxies should be infinitely hot and bright. Since neither of these are observable phenomena in the universe, [323] claims that temperature must be invariant with the speed of the inertial reference frame.

The caveat in the aforementioned argument lies in the fact that if the energy density in Einstein’s field equation is positive capturing the effect of an expanding universe with distant galaxies moving at very high speeds, then the associated negative pressure captured by the Riemann curvature tensor results in an accelerated expansion of the universe. Thus, even though the temperature invariance
assumption will hold for a noninertial observer outside the system, that does not imply that temperature is invariant if the system observed is accelerating and/or the system is subjected to a strong gravitational field.

Alternatively, if we assume that the expansion speed is constant and use Planck’s formula [339] for characterizing the intensity of light as a function of frequency and temperature, then there is a key difference between the frequency shift of light due to the relativistic Doppler effect and the shift of light due to the change in temperature. Specifically, the relativistic Doppler effect would account for changes in the frequency of light due to relativistic motion between a moving light source and an observer, where time dilation results in a larger number of oscillations in time when compared to the number of oscillations in time of the moving inertial frame.

Even if we separate the light intensity change due to the relativistic Doppler effect from that of the temperature, then the shift attributed to the temperature is a product of both the temperature \( T \) and the Boltzmann constant \( k \). Thus, since entropy is no longer invariant in this formulation, the product \( kT \), and not just the Boltzmann constant \( k \), can relativistically scale leading to a similar indeterminacy as in (55). This can produce the observed effect we see as distant galaxies move away in our expanding universe without relativistic temperature invariance necessarily holding.

The authors in [333] also arrive at similar conclusions as in [323] using a set-theoretic definition of temperature involving a bijective (i.e., one-to-one and onto) continuous mapping of a manifold \( \mathcal{H} \) defined on a simply connected subset \( \mathcal{I} \) of the reals \( \mathbb{R} \). Specifically, it is shown that \( \mathcal{H} \) is topologically equivalent to \( \mathbb{R} \) and contains a countably dense and unbounded subset \( \mathcal{F}_{th} \) of all the thermometric fixed points contained in \( \mathcal{H} \). Thermometric fixed points characterize observable qualitative properties of a physical system state such as the boiling and freezing points of water at atmospheric pressure and provide a countable dense subset in \( \mathcal{H} \) having a topological structure involving a closed set homeomorphic to the real line. The properties of the set \( \mathcal{F}_{th} \) and its relation to the manifold \( \mathcal{H} \) are characterized by a set of postulates satisfying empirical observations. Furthermore, the mapping of the manifold \( \mathcal{H} \) induces a homeomorphism on the set of positive real numbers characterizing the Kelvin temperature scale.

Using the assertion that thermometric fixed points have the same behavior in all inertial frames—a statement first proffered in [323]—the authors in [333] claim that temperature must be a Lorentz invariant. Specifically, the authors in [333] argue that water boiling in a system at rest cannot “simultaneously” not be boiling if observed from an inertially moving system. Hence, they conclude that every thermometric fixed point has to “correspond to the same hotness level” regardless of which inertial frame is used to observe the event.

What has been overlooked here, however, is that relativity is a theory of quantitative measurements, and how motion affects these measurements, and not what we perceive through observations. In other words, what is perceived through observation might not necessarily correspond to what is measured. This statement is equally valid for stationary systems. For example, in Rayleigh flow the stagnation temperature is a variable, and hence, in regions wherein the flow accelerates faster than heat is added, the addition of heat to the system causes the system temperature to decrease. This is known as the Rayleigh effect and is a clear example of what is deduced through observation does not necessarily correspond to what is measured. Furthermore, since simultaneity is not independent of the frame of reference used to describe the water boiling event, the Lorentz transformation of time plays a key role in the observed process. In addition, even though the pressure controlling the boiling point of water is Lorentz invariant, the volume is not.

To further elucidate the above discussion, we turn our attention to a perfect (i.e., ideal) monatomic gas. The relation between the pressure, volume, and temperature of such a gas follows from the perfect (ideal) gas law and is given by

\[
pV = NkT,
\]  
(63)
where $p$ is the pressure of the gas, $V$ is the volume of the gas, $N$ is the number of molecules present in the given volume $V$, $k$ is the Boltzmann constant relating the temperature and kinetic energy in the gas, and $T$ is the absolute temperature of the gas. Recall that the Boltzmann constant $k$ is given by the ratio of the gas law constant $R$ for one mole of the gas to Avogadro’s number $N_A$ quantifying the number of molecules in a mole, that is,

$$k = \frac{R}{N_A}. \quad (64)$$

Now, relating the number of moles (i.e., the number density) to the number of molecules $N$ in a given volume and the number of molecules in a mole $N_A$ by $n = N/N_A$, (63) can be equivalently written as

$$pV = nRT. \quad (65)$$

It is important to stress here that a perfect monatomic gas is an idealization and provides a useful abstraction in physics that has been used to develop the ideal gas temperature scale through (63). This temperature scale is totally arbitrary and has been concomitantly associated with the definition and measurement of temperature for historical and practical reasons. Furthermore, the molecular Equation (63) is derived from statistical mechanics and does not provide a phenomenological definition of temperature as used in mainstream physics, chemistry, and engineering predicated on macroscopic sensors (e.g., thermometers). Invoking probability theory and the badly understood concepts of statistical mechanics in an attempt to define the already enigmatic physical concept of temperature has led to a lack of a satisfactory definition of temperature in the literature [340].

Since by the principle of relativity the ideal gas law must hold for all reference frames, it follows from (63) that

$$p'V' = NK'T' \quad (66)$$

and

$$pV = NkT, \quad (67)$$

where the number of molecules $N$ present in a given volume is relativistically invariant for every inertial frame. Now, using (56) and (57) it follows from (66) and (67) that

$$p'V' = pV\sqrt{1 - (u/c)^2}$$
$$= NKT\sqrt{1 - (u/c)^2}$$
$$= NK'T', \quad (68)$$

which yields

$$k'T' = kT\sqrt{1 - (u/c)^2}. \quad (69)$$

Note that (69) has the same form as (45) confirming that the relativistic transformation of the change in the product of entropy and temperature $dS'T'$, which can be characterized by (69), is equivalent to the relativistic change in heat $dQ'$. Next, using (55) and (62) it follows from (69) that

$$k' = f(\beta)k, \quad (70)$$

which shows that the Boltzmann constant is not invariant with speed. Furthermore, since the Boltzmann constant $k$ is related to the gas law constant $R$ through the number of molecules $N_A$ in a mole and $N_A$ is a relativistic invariant, it follows from (64) and (70) that

$$R' = f(\beta)R. \quad (71)$$
If we can ascertain that temperature is a relativistic invariant, then \( T' = T \), and hence, \( g(\beta) = 1 \) and \( f(\beta) = \sqrt{1 - (u/c)^2} \). In this case, it follows from (70) and (71) that
\[
k' = k\sqrt{1 - (u/c)^2}
\]
and
\[
R' = R\sqrt{1 - (u/c)^2}.
\]

This would imply that the relativistic change in entropy transforms exactly as the relativistic change in time. However, in the absence of experimental verification of relativistic temperature invariance given the present state of our scientific and technological capabilities, as well as the indeterminacy condition manifested through the mathematical analysis of the relativistic second law leading to (55), the assertion of relativistic temperature invariance is unverifiable [341].

Both (70) and (72) have far reaching consequences in physics and cosmology. The Boltzmann constant relates the translational kinetic energy of the molecules of an ideal gas at temperature \( T \) by [342–344].
\[
E_{KE} = \frac{3}{2} N k T.
\]

In classical physics, \( k \) is a universal constant that is independent of the chemical composition, mass of the molecules, substance phase, etc. Relativistic thermodynamics, however, asserts that even though the Boltzmann constant can be assumed to be independent of such physical variables, it is not Lorentz invariant with respect to the relative velocities of inertially moving bodies.

Finally, we close this section by showing that when the relative velocity \( u \) between the two reference frames \( F \) and \( F' \) approaches the speed of light \( c \), the rate of change in entropy increase of the moving system decreases as the system’s speed increases through space. Specifically, note that it follows from (61) that
\[
\frac{dS'}{dt'} = \frac{dS}{dt} f(\beta),
\]
which, using the relativistic equation connecting differential time intervals in the two reference frames \( F' \) and \( F \) between neighboring events occurring at neighboring points in space given by [9]
\[
\frac{dt'}{dt} = \frac{1 - (v/c)^2}{\sqrt{1 - (v/c)^2}} \frac{dx}{dt} = \frac{1 - (v/c)^2 u_x}{\sqrt{1 - (v/c)^2}},
\]
where \( u_x \triangleq \frac{dx}{dt} \), it follows from (75) that
\[
\frac{dS'}{dt'} = \left[ \frac{\sqrt{1 - (u/c)^2}}{1 - (u/c)^2 u_x} \right] f(\beta) \frac{dS}{dt}.
\]

Thus,
\[
\lim_{u \to c} \frac{dS'}{dt'} = 0,
\]
which shows that motion affects the rate of entropy increase. In particular, the change in system entropy is observer dependent and is fastest when the system is at rest relative to an observer. Hence, the change in entropy of an inertially moving system relative to a stationary observer decreases leading to an entropy contraction. Conversely, an observer in a nonproper reference frame detects a dilated change in entropy.
16. General Relativity and Thermodynamics

The extension of thermodynamics to general relativity is far more complex and subtle than the unification of special relativity with thermodynamics presented in Section 14. The most complete formulation attempt of general relativity and thermodynamics was developed by Tolman [316]. However, Tolman’s general relativistic thermodynamics, along with its incremental extensions in the literature, are restricted to equilibrium thermodynamic systems and are almost all exclusively predicated on the archaic assertion that entropy is a Lorentz invariant. Furthermore, the theory is based on homogeneous models for the distribution of material leading to time-invariant tensor densities in different reference frames. These assumptions along with the unsound assertion of the invariance of entropy and the demise of the gas law and Boltzmann constants as universal constants places these formulations of general relativistic thermodynamics in serious peril.

To viably extend thermodynamics to general relativity and address the universality of thermodynamics to cosmology, the dynamical system framework formulation of thermodynamics developed in [9] needs to be merged with general relativity theory, wherein the tensor currents in the Einstein field equation are explicitly dependent on time for every noninertial reference frame. This will lead to time-varying light cone observables further leading to spatial quantum fluctuations. Furthermore, since the thermodynamic variables of entropy, temperature, heat, and work are global quantities, and matter (i.e., energy and momentum) distorts (i.e., curves) spacetime in the presence of a gravitational field, spacetime curvature (i.e., confinement) can violate global conservation laws leading to singularities when integrating tensor densities to obtain these global thermodynamic variables.

The extension of the dynamical system framework of thermodynamics to general relativity is beyond the scope of this paper as it requires several extensions of the developed system thermodynamic framework as well as additional machinery using tensor calculus. With the aforementioned complexities notwithstanding, the notions of a thermal equilibrium in the presence of a gravitational field, the direction of heat flow in the presence of momentum transfer, and the continued progression of an isolated thermodynamic processes that does not (asymptotically or in finite time) converge to an equipartitioned energy state corresponding to a state of maximum entropy all need to be reviewed.

Even the method of measurement for assessing the existence of a thermal equilibrium between moving inertial systems can also be brought into question. If heat is exchanged between relatively moving inertial systems, then by the energy-momentum conservation principle there must also exist a momentum exchange between the systems. In this case, Does heat necessarily flow in the direction of lower temperatures?

A strong gravitational field brings the notion of a uniform temperature distribution (i.e., Axiom i of [3]) as a necessary condition for thermal equilibrium into question as well. Specifically, in the presence of a strong gravitational field, a temperature gradient is necessary for preventing the flow of heat from a higher to a lower gravitational potential because of the equivalence between mass and energy. In this case, a uniform temperature distribution would no longer be a necessary condition for the establishment of a thermal equilibrium. In other words, in general relativity a temperature gradient is necessary at thermal equilibrium for preventing heat flow from a region of a higher gravitational potential to a region of a lower gravitational potential.

Another important consideration is establishing a physically accurate kinetic equation for the description of heat transfer with finite thermal propagation speeds in the spacetime continuum. It is well known that a parabolic wave diffusion equation for characterizing thermal transport allows for the physically flawed prediction of an infinite wave speed of heat conduction between bodies in direct thermal contact [345–349]. Since parabolic heat transfer equations in the temperature field can be formulated as a consequence of the first and second laws of thermodynamics [346,350], they confirm a thermodynamical prediction of infinite thermal propagation in heat conduction.

More specifically, Fourier’s heat diffusion equation describes the energy flow distance from the origin of a fixed inertial frame increasing proportionally with the square root of the proper time resulting in a proper energy transport velocity increasing inversely proportional with the square
root of the proper time [345] (p. 146). Over short time scales, this leads to a contradiction between thermodynamics and relativity theory. In particular, the fact that relativity theory asserts that the speed of light is an upper limit for all possible velocities between material systems, superluminal heat propagation speeds contradicts the establishment of a virtually instantaneous thermal equilibrium between bodies in short proximity as predicted by classical thermodynamics [346].

In an attempt to develop a consistent description of heat conduction with finite speed of heat propagation, several heat transport models have been introduced in the literature [347–349,351,352]. The most common hyperbolic heat conduction model yielding a finite speed of thermal propagation was developed by Cattaneo [353] and Vernotte [354]. The Cattaneo-Vernotte transport law involves a modification of the classical Fourier law that transforms the parabolic heat equation into a damped wave equation by including a regularization term in the equation with a heat flux relaxation constant.

Even though several hyperbolic heat conduction models have been developed in the literature to capture finite thermal wave speeds, many of these models display the unphysical behavior that, in an adiabatically isolated system, heat can flow from regions of lower to higher temperatures over finite time intervals contradicting the second law of thermodynamics. In addition, some models give temperature predictions below absolute zero. This motivated authors to develop alternative hyperbolic heat conduction equations consistent with rational extended thermodynamics based on nonphysical interpretations of absolute temperature and entropy notions [109,355]. A notable exception to this is the model developed in [346], which gives a classical thermodynamically consistent description of heat conduction with finite speed of heat conduction.

Finally, it is also not clear whether an infinite and endlessly expanding universe governed by the theory of general relativity has a final equilibrium state corresponding to a state of maximum entropy. The energy density tensor in Einstein’s field equation is only covariantly conserved when changes in the curvature of the spacetime continuum need to be addressed since it does not account for gravitational energy. It follows from relativistic mechanics that the total proper energy (i.e., rest energy) in an isolated system need not be constant. Thus, cosmological models can undergo cyclical expansions and contractions as long as there exists a sufficient amount of external energy to continue this cyclical morphing while maintaining an increase in both energy and entropy, with the later never reaching a maximum value. In other words, cosmological models can be constructed that do not possess a limitation on the total proper energy, and hence, would also not possess an upper bound on the system entropy [316].

17. Conclusions

In this paper, we traced the long and tortuous history of thermodynamics from its classical to its postmodern forms. Furthermore, we provided a tutorial exposition of the general systems theory framework for thermodynamics developed in [3,9], which attempts to harmonize thermodynamics with classical mechanics. As outlined in the paper, the theory of thermodynamics followed two conceptually rather different schools of thought, namely, the macroscopic point of view versus the microscopic point of view.

The microscopic point of view of thermodynamics was first established by Maxwell [356] and further developed by Boltzmann [61] by reinterpreting thermodynamic systems in terms of molecules or atoms. However, since the microscopic states of thermodynamic systems involve a large number of similar molecules, the laws of classical mechanics were reformulated so that even though individual atoms are assumed to obey the laws of Newtonian mechanics, the statistical nature of the velocity distribution of the system particles corresponds to the thermodynamic properties of all the atoms together. This resulted in the birth of statistical mechanics. The laws of mechanics, however, as established by Poincaré [20], show that every isolated mechanical system will return arbitrarily close to its initial state infinitely often. Hence, entropy must undergo cyclic changes and thus cannot monotonically increase. This is known as the recurrence paradox or Loschmidt’s paradox.
In statistical thermodynamics the recurrence paradox is resolved by asserting that, in principle, the entropy of an isolated system can sometimes decrease. However, the probability of this happening, when computed, is incredibly small. Thus, statistical thermodynamics stipulates that the direction in which system transformations occur is determined by the laws of probability, and hence, they result in a more probable state corresponding to a higher system entropy. However, unlike classical thermodynamics, in statistical thermodynamics it is not absolutely certain that entropy increases in every system transformation. Hence, thermodynamics based on statistical mechanics gives the most probable course of system evolution and not the only possible one, and thus heat flows in the direction of lower temperature with only statistical certainty and not absolute certainty. Nevertheless, general arguments exploiting system fluctuations in a systematic way [23] seem to show that it is impossible, even in principle, to violate the second law of thermodynamics.

The newly developed dynamical system notions of entropy proposed in [3,9] as well as in [111–113] involving an analytical description of an objective property of matter can potentially offer a conceptual advantage over the subjective quantum expressions for entropy proposed in the literature (e.g., Daróczy entropy, Hartley entropy, Rényi entropy, von Neumann entropy, infinite-norm entropy) involving a measure of information. An even more important benefit of the dynamical systems representation of thermodynamics is the potential for developing a unified classical and quantum theory that encompasses both mechanics and thermodynamics without the need for statistical (subjective or informational) probabilities. This can pave the way for designing future micromechanical machines [357,358].

There is no doubt that thermodynamics plays a fundamental role in cosmology, physics, chemistry, biology, engineering, and the information sciences as its universal laws represent the condiciones sine quibus non in understanding Nature. Thermodynamics, however, has had an impact far beyond the physical, chemical, biological, and engineering sciences; its outreach has reverberated across disciplines, with philosophers to mathematicians, anthropologists to physicians and physiologists, economists to sociologists and humanists, writers to artists, and apocalypticists to theologians each providing their own formulations, interpretations, predictions, speculations, and fantasies on the ramifications of its supreme laws governing Nature, all of which have added to the inconsistencies, misunderstandings, controversies, confusions, and paradoxes of this aporetic yet profound subject.

The fact that almost everyone has given their own opinion regarding how thermodynamics relates to their own field of specialization has given thermodynamics an anthropomorphic character. In that regard, Bridgman [69] (p. 3) writes “It must be admitted, I think, that the laws of thermodynamics have a different feel from most other laws of the physicist. There is something more palpably verbal about them—they smell more of their human origin”. The second law has been applied to the understanding of human society as well as highlighting the conflict between evolutionists, who claim that human society is constantly ascending, and degradationists, who believe that the globalization processes in human societies will lead to equalization (i.e., equipartitioning), social degradation, and the lack of further progress—a transition towards an idiocratic state of human society in analogy to the thermodynamic heat death [359].

Thermodynamics and entropy have also found their way into economics, wherein human society is viewed as a superorganism with the global economy serving as its substrate [360,361]. In particular, the complex, large-scale material structure of civilization involving global transportation and telecommunication networks, water supplies, power plants, electric grids, computer networks, massive goods and services, etc., is regarded as the substratum of human society, and its economy is driven by ingesting free energy (hydrocarbons, ores, fossil fuels, etc.), material, and resources, and secreting them in a degraded form of heat and diluted minerals resulting in an increase of entropy.

Connections between free energy and wealth, wealth and money, and thermodynamic depreciation reflected in the value of money over time via purchasing power loss and negative interest rates have also been addressed in the literature. Soddy [360] maintained that real wealth is subject to the second law of thermodynamics; namely, as entropy increases, real wealth deteriorates. In particular, there are limits
to growth given the finite resources of our planet, and there exists a balance between the economy and our ecosphere. In other words, the rate of production of goods has to be delimited by the sustainable capacity of the global environment.

The economic system involves a combination of components that interact between constituent parts of a larger dynamical system, and these interactions cause the degradation of high-grade energy (free energy) into low-grade energy and, consequently, an increase in entropy. The vanishing supplies of petroleum and metal ores, deforestation, polluted oceans, cropland degradation by erosion and urbanization, and global warming, are all limiting factors in economics. As we continue to demand exponential economic growth in the face of a geometrically increasing population that exceeds the environment’s sustainability, this will inevitably result in a finite-time bifurcation leading to a catastrophic collapse of the ecosphere.

Even though such analogies to thermodynamics abound in the literature for all of the aforementioned disciplines, they are almost all formulated in a pseudoscientific language and a fragile mathematical structure. The metabasis of one discipline into another can lead to ambiguities, dilemmas, and contradictions; this is especially true when concepts from a scientific discipline are applied to a nonscientific discipline by analogy. Aristotle (384–322 B.C.) was the first to caution against the transition or metabasis—μεταβάσις—of one science into another, and that such a transition should be viewed incisively.

Perhaps the most famous example of this is Malthus’ 1798 essay On the principle of population as it affects the future improvement of society. His eschatological missive of a geometrical population growth within an arithmetic environmental growing capacity reverberates among economists from Malthus’ time to the present. Heilbroner [362] writes: “In one staggering intellectual blow Malthus undid all the roseate hopes of an age oriented toward self-satisfaction and a comfortable vista of progress”. After reading Malthus, Thomas Carlyle refers to the field of economics as “the dismal science”. Of course, Malthus’ argument is hypothetical, and his conclusion importunate; such assertions, however, are neither provable nor refutable.

Given the universality of thermodynamics and the absolute reign of its supreme laws over Nature, it is not surprising that numerous authors ranging from scientists to divinests have laid claim to this unparalleled subject. This has, unfortunately, resulted in an extensive amount of confusing, inconsistent, and paradoxical formulations of thermodynamics accompanied by a vague and ambiguous lexicon giving it its anthropomorphic character. However, if thermodynamics is formulated as a part of mathematical dynamical systems theory—the prevailing language of science, then thermodynamics is the irrefutable pathway to uncovering the deepest secrets of the universe.

Acknowledgments: This research was supported in part by the Air Force Office of Scientific Research under Grant FA9550-16-1-0100.

Conflicts of Interest: The author declares no conflict of interest.

References and Notes


8. Complexity here refers to the quality of a system wherein interacting subsystems self-organize to form hierarchical evolving structures exhibiting emergent system properties.


14. Many natural philosophers have associated this ravaging irrecoverability in connection to the second law of thermodynamics with an eschatological terminus of the universe. Namely, the creation of a certain degree of life and order in the universe is inevitably coupled with an even greater degree of death and disorder. A convincing proof of this bold claim has, however, never been given.

15. The earliest perception of irreversibility of nature and the universe along with time’s arrow was postulated by the ancient Greek philosopher Herakleitos (~535–~475 B.C.). Herakleitos’ profound statements, *Everything is in a state of flux and nothing is stationary and Man cannot step into the same river twice, because neither the man nor the river is the same*, created the foundation for all other speculation on metaphysics and physics. The idea that the universe is in constant change and that there is an underlying order to this change—the *Logos*—postulates the very existence of entropy as a physical property of matter permeating the whole of nature and the universe.


24. The theory of classical thermodynamics has also been developed over the last one and a half centuries by many other researchers. Notable contributions include the work of Maxwell, Rankine, Clapeyron, Bridgman, Kestin, Meixner, and Gile.


26. A *perpetuum mobile* of the second kind is a cyclic device that would continuously extract heat from the environment and completely convert it into mechanical work. Since such a machine would not create energy, it would not violate the first law of thermodynamics. In contrast, a machine that creates its own energy and thus violates the first law is called a *perpetuum mobile* of the first kind.

27. Carnot never used the terms reversible and irreversible cycles, but rather cycles that are performed in a inverse direction and order [28] (p. 11). The term reversible was first introduced by Kelvin [29] wherein the cycle can be run backwards.


30. After Carnot’s death, several articles were discovered wherein he had expressed doubt about the caloric theory of heat (i.e., the conservation of heat). However, these articles were not published until the late 1870’s, and as such, did not influence Clausius in rejecting the caloric theory of heat and deriving Carnot’s results using the energy equivalence principle of Mayer and Joule.

31. Μὲν οὗν φησιν εἷναι τὸ πᾶν διαιρετὸν ἀδιαίρετον, γενητὸν ἀγένητον, θνητὸν ἀθάνατον, λὸγον αἰῶνα, πατέρα υίον, ...ἐστίν ἕν πάντα εἷναι.

32. Φύσις ουδενός εστίν εόντων αλλά μόνον μίξις τε, διάλλαξίς τε μιγέντων εστί, φύσις δ’ επί τους ονομάζεται ανθρώπων—There is no genesis with regards to any of the things in nature but rather a blending and alteration of the mixed elements; man, however, uses the word ‘nature’ to name these events.


34. Clausius, R. "Über verschiedene füre die Anwendung bequeme Formen der Haubtgleichungen der mechanischen wärmetheorie. Viertelsjahrschrift der Naturforschenden Gesellschaft (Zürich) 1865, 10, 1–59. (In German)

35. Clausius, R. Abhandlungen Über Die Mechanische Wärme-Theorie; Vieweg and Sohn: Braunschweig, Germany, 1867. (In German)

36. Clausius, R. Mechanische Wärmetheorie; Vieweg and Sohn: Braunschweig, Germany, 1876. (In German)

37. Clausius succinctly expressed the first and second laws of thermodynamics as: “Die energie der Welt ist konstant und die entropie der Welt strebt einem maximum zu”. Namely, the energy of the Universe is constant and the entropy of the Universe tends to a maximum.

38. Many conservation laws are a special case of Nöether’s theorem which states that for every one-parameter group of diffeomorphisms defined on an abstract geometrical space (e.g., configuration manifolds, Minkowski space, Riemannian space) of a Hamiltonian dynamical system that preserves a Hamiltonian function, there exist first integrals of motion. In other words, the algebra of the group is the set of all Hamiltonian systems whose Hamiltonian functions are the first integrals of motion of the original Hamiltonian system.


40. A world manifold is a four-dimensional orientable, noncompact, parallelizable manifold that admits a semi-Riemannian metric and a spin structure. Gravitation theories are formulated on tensor bundles that admit canonical horizontal prolongations on a vector field defined on a world manifold. These prolongations are generators of covariant transformations whose vector field components play the role of gauge parameters. Hence, in general relativity the energy-momentum flow collapses to a superpotential of a world vector field defined on a world manifold admitting gauge parameters.


43. In the case of thermodynamic systems with positive absolute temperatures, Kelvin’s postulate can be shown to be equivalent to Clausius’ postulate. However, many textbooks erroneously show this equivalence without the assumption of positive absolute temperatures. Physical systems possessing a small number of energy levels with negative absolute temperatures are discussed in [44–48].


49. Κόσμον (τόνδε), τὸν αὐτὸν ἁπάντων, οὔτε τις θεῶν, οὔτε ἀνθρώπων ἐποίησεν, ἀλλ´ ᾖν ἀεὶ καὶ ἔστιν καὶ ἔσται πῦρ ἀείζωον, ἁπτόμενον μέτρα καὶ ἀποσβεννύμενον μέτρα.

50. Είναι τε ώσπερ γενέσεις κόσμου, οὕτω καί αυξήσεις καί φθίσεις καί φθοράς, κατά τινά ανάγκην.

51. Planck, M. Vorlesungen Über Thermodynamik; Veit: Leipzig, Germany, 1897. (In German)

53. Truesdell [54] (p. 328) characterizes the work as a “gloomy murk”, whereas Khinchin [55] (p. 142) declares it an “aggregate of logical mathematical errors superimposed on a general confusion in the definition of the basic quantities”.


57. Gibbs’ principle is weaker than Clausius’ principle leading to the second law involving entropy increase since it holds for the more restrictive case of isolated systems.


60. Carathéodory’s definition of an adiabatic process is nonstandard and involves transformations that take place while the system remains in an adiabatic container; this allowed him to avoid introducing heat as a primitive variable. For details see [58,59].


62. The number of distinct microstates $W$ can also be regarded as the number of solutions of the Schrödinger equation for the system giving a particular energy distribution. The Schrödinger wave equation describes how a quantum state of a system evolves over time. The solution of the equation characterizes a wave function whose wavelength is related to the system momentum and frequency is related to the system energy. Unlike Planck’s discrete quantum transition theory of energy when light interacts with matter, Schrödinger’s quantum theory stipulates that quantum transition involves vibrational changes from one form to another; and these vibrational changes are continuous in space and time.


72. The phrase arrow of time was coined by Eddington in his book The Nature of the Physical World [71] and connotes the one-way direction of entropy increase educed from the second law of thermodynamics. Other phrases include the thermodynamic arrow and the entropic arrow of time. Long before Eddington, however, philosophers and scientists addressed deep questions about time and its direction.

73. Parmenides (~515–~450 b.c.) maintained that there is neither time nor motion. His pupil Zeno of Elea (~490–~430 b.c.) constructed four paradoxes—the dichotomy, the Achilles, the flying arrow, and the stadium—to prove that motion is impossible. His logic was “immeasurably subtle and profound” and even though infinitesimal calculus provides a tool that explains Zeno’s paradoxes, the paradoxes stand at the intersection of reality and our perception of it; and they remain at the cutting edge of our understanding of space, time, and spacetime [74].


75. It is interesting to note that, despite his steadfast belief in change, Herakleitos embraced the concept of eternity as opposed to Parmenides’ endless duration concept.
76. Perhaps a better expression here is the geodesic arrow of time, since, as Einstein’s theory of relativity shows, time and space are intricately coupled, and hence one cannot curve space without involving time as well. Thus, time has a shape that goes along with its directionality.
84. Plato (~428–~348 B.C.) writes that time was created as an image of the eternal. While time is everlasting, time is the outcome of change (motion) in the universe. And as night and day and month and the like are all part of time, without the physical universe time ceases to exist. Thus, the creation of the universe has spawned the arrow of time—Χρόνον τε γενέσθαι ειδόν τοῦ οὐδού. Κύκλωσαν μὲν αἰεί μένειν, τὴν δὲ τοῦ οὐρανοῦ φοράν χρόνον εἶναι: καὶ γὰρ νύκτα καὶ ημέρα καὶ μῆνα καὶ τὰ τοιαῦτα πάντα χρόνου μέρη εἶναι. Διότερρα δὲν τῆς τοῦ κόσμου ύπόσεως ύπόπτες οὐ εἶναι χρόνον ἄει γὰρ ὑπάρχειν αὐτῶ καὶ χρόνον εἶναι.
85. In statistical thermodynamics the arrow of time is viewed as a consequence of high system dimensionality and randomness. However, since in statistical thermodynamics it is not absolutely certain that entropy increases in every dynamical process, the direction of time, as determined by entropy increase, has only statistical certainty and not an absolute certainty. Hence, it cannot be concluded from statistical thermodynamics that time has a unique direction of flow.
86. There is an exception to this statement involving the laws of physics describing weak nuclear force interactions in Yang-Mills quantum fields [87]. In particular, in certain experimental situations involving high-energy atomic and subatomic collisions, meson particles (K-mesons and B-mesons) exhibit time-reversal asymmetry [88]. However, under a combined transformation involving charge conjugation C, which replaces the particles with their antiparticles, parity P, which inverts the particles’ positions through the origin, and a time-reversal involution R, which replaces t with −t, the particles’ behavior is CPT-invariant. For details see [88].
90. Conversely, one can also find many authors that maintain that the second law of thermodynamics has nothing to do with irreversibility or the arrow of time [91–93]; these authors largely maintain that thermodynamic irreversibility and the absence of a temporal orientation of the rest of the laws of physics are disjoint notions. This is due to the fact that classical thermodynamics is riddled with many logical and mathematical inconsistencies with carelessly defined notation and terms. And more importantly, with the notable exception of [3], a dynamical systems foundation of thermodynamics is nonexistent in the literature.
94. Prigogine, I. From Being to Becoming; Freeman: San Francisco, CA, USA, 1980.
95. The Higgs boson is an elementary particle (i.e., a particle with an unknown substructure) containing matter (particle mass) and radiation (emission or transmission of energy), and is the finest quantum constituent of the Higgs field.
110. A key exception here is irreversible mechanothermodynamics [111–113] involving irreversible damage of complex system states discussed in Section 6.
140. In relativistic physics, an event horizon is a boundary delineating the set of points in spacetime beyond which events cannot affect an outside observer. In the present context, it refers to the boundary beyond which events cannot escape the black hole’s gravitational field.
161. The Hellenistic period (323–31 B.C.) spawned the scientific revolution leading to today’s scientific method and scientific technology, including much of modern science and mathematics in its present formulation. Hellenistic scientists, which included Archimedes, Euclid, Eratosthenes, Eudoxus, Ktesibios, Philo, Apollonios and many others, were the first to use abstract mathematical models and attach them to the physical world. More importantly, using abstract thought and rigorous mathematics (Euclidean geometry, real numbers, limits, definite integrals) these “modern minds in ancient bodies” were able to deduce complex solutions to practical problems and provide a deep understanding of nature. In his *Forgotten Revolution* [162] Russo convincingly argues that Hellenistic scientists were not just forerunners or anticipators of modern
science and mathematics, but rather the true fathers of these disciplines. He goes on to show how science was born in the Hellenistic world and why it had to be reborn.

As in the case of the origins of much of modern science and mathematics, modern engineering can also be traced back to ancient Greece. Technological marvels included Ktesibios’ pneumatics, Heron’s automata, and arguably the greatest fundamental mechanical invention of all time—the Antikythera mechanism. The Antikythera mechanism, most likely inspired by Archimedes, was built around 76 B.C. and was a device for calculating the motions of the stars and planets, as well as for keeping time and calendar. This first analog computer involving a complex array of meshing gears was a quintessential hybrid dynamical system that unequivocally shows the singular sophistication, capabilities, and imagination of the ancient Greeks, and dispenses with the Western myth that the ancient Greeks developed mathematics but were incapable of creating scientific theories and scientific technology.


163. ὡς γὰρ ὁ ἥλιος εἰς ἑαυτὸν ἐπιστρέφει τὰ μέρη ἐξ ὧν συνέστηκε, καὶ ἡ γῆ (in Ploutarkhos, De facie quae in orbe lunae apparat, 924E).

164. In his treatise on The Method of Mechanical Theorems Archimedes (287–212 B.C.) established the foundations of integral calculus using infinitesimals, as well as the foundations of mathematical mechanics. In addition, in one of his problems he constructed the tangent at any given point for a spiral, establishing the origins of differential calculus [165] (p. 32).


169. Lagrange, J.L. Méchanique Analitique; Desaint: Paris, France, 1788. (In French)


175. Lyapunov, A.M. The General Problem of the Stability of Motion; Kharkov Mathematical Society: Kharkov, Russia, 1892.


Entropy 2017, 19, 621


Entropy comes from the Greek word Εντροπία (εν and τροπή) for outward transformation connoting evolution or complexification and is the literal antonym of entropy (εντροπία—εν and τροπή), signifying an inward transformation connoting devolution or decomplexification. The word entropy was proposed by Clausius for its phonetic similarity to energy with the additional connotation reflecting change (τροπή).

In the terminology of [70], state irreversibility is referred to as time-reversal non-invariance. However, since the term time-reversal is not meant literally (that is, we consider dynamical systems whose trajectory reversal is or is not allowed and not a reversal of time itself), state reversibility is a more appropriate expression. And in that regard, a more appropriate expression for the arrow of time is the deterioration of time signifying irrecoverable system changes.

Here we use the term free energy to denote either the Helmholtz free energy or the Gibbs free energy depending on context. Recall that the Helmholtz free energy \( F = U - TS \) is the maximal amount of work a system can perform at a constant volume and temperature, whereas the Gibbs free energy \( G = U - TS - pV \) is the maximal amount of work a system can perform at constant pressure and temperature. Hence, if pressure gradients can perform useful work and actuate organization (i.e., hurricanes, shock waves, tornados), then the Helmholtz free energy is the most relevant free energy. Alternatively, if pressure is constant and changes in volume need to be accounted for, then the Gibbs free energy is the relevant free energy.

It is important to stress here that even though energy is conserved and cannot be created or destroyed, free energy or, more descriptively, extractable useful work energy, can be destroyed. Every dynamical process in nature results in the destruction (i.e., degradation) of free energy.

The external source of energy of almost all life on Earth is principally supplied by solar radiation. An interesting exception to this are the deep ocean volcanic vent ecosystems that derive their energy from endogenous heat sources due to radioactive material and chemical gradients emanating from volcanic activities on the ocean floor.


Photon energy from the Sun is generated as gamma rays that are produced by thermonuclear reactions (i.e., fusion) at the center of the Sun and distributed among billions of photons through the Sun’s photosphere.
Planck’s work on thermodynamics and black-body radiation led him to formulate the foundations of quantum theory, wherein electromagnetic energy is viewed as discrete amounts of energy known as quanta or photons. The Planck quantum formula relates the energy of each photon \( E \) to the frequency of radiation \( \nu \) as 
\[
E = h \nu,
\]
where \( h \) is the Planck constant.

When photons are absorbed by the Earth, they induce electromagnetic transmissions in matched energy absorber bands leading to photochemistry decay and fluorescence, phosphorescence, and infrared emissions.

Here we are assuming that the average temperature of the Earth is constant, and hence, the amount of energy delivered by solar photons to the Earth is equal to the amount of energy radiated by infrared photons from the Earth. If this were not the case, then the internal energy of the Earth \( U_{\text{Earth}} \) would increase resulting in a rise of the Earth’s average temperature.

Shannon introduces a constant \( K \) into his information entropy formula in only two places stating that \( K \) can be introduced as a matter of convenience and can be used to attach a choice of a unit measure.

It follows from general relativity that pressure generates gravity. Thus, repulsive gravity involves an internal gravitational field acting within space impregnated by negative pressure. This negative pressure generates a repulsive gravitational field that acts within space [226,227].

The brain has the most abounding energy metabolism in the human body. Specifically, in a resting awake state, it requires approximately 20% of the total oxygen supplied by the respiratory system and 25% of the total body glucose [232,233]. Cerebral glucose metabolism is critical for neural activity. This can be seen in hypoglycemic individuals with diminished glucose metabolism experiencing impaired cognitive function.

Every living organism that is deprived of oxygen consumption and glycolysis dies within a few minutes as it can no longer produce heat and release entropy to its environment. In other words, its homeostatic state is destroyed.

Aoki introduces the min-max principle of entropy production with time in aquatic communities. [235]


242. In addition to the central nervous system, entropy production over the course of a living mammalian organism is directly influenced by dissipation of energy and mass to the environment. However, the change in entropy due to mass exchange with the environment is negligible (approximately 2%), with the predominant part of entropy production attributed to heat loss due to radiation and water evaporation.


244. A fractal is a set of points having a detailed structure that is visible on arbitrarily small scales and exhibits repeating patterns over multiple measurement scales. A fractal dimension is an index measure of complexity capturing changes in fractal patterns as a function of scale.

245. There are exceptions to this pattern. For example, most cancers including gliomas, central nervous system lymphomas, and pituitary lesions are hypermetabolic resulting in a high-rate of glycolysis leading to an increased entropy production and fractal dimension [236]. This accession of entropy production, however, is deleterious to the host organism as it results in an over expenditure of the energy substratum; see [236].


254. Mechanistic models in this context are models that are predicated on dynamical systems theory wherein an internal state model is used to describe dynamic linking between phenotypic states using biological and physiological laws and system interconnections.


260. Here we adopt a scientific (i.e., neuroscience) perspective of consciousness and not a philosophical one. There have been numerous speculative theories of human consciousness, with prominent metaphysical theories going back to ancient Greece and India. For example, Heraclitus proposed that there exists a single, all-powerful, divine consciousness that controls all things in Nature and that ultimate wisdom is reached when one achieves a fundamental understanding of the universal laws that govern all things and all forces in the universe—Εἶναι γὰρ ἓν τὸ σοφόν, ἐπίρτασθαι γνώμην, ὁτέῃ ἐκφύσηνεν πάντα ἅκα πάνων. In Hinduism, Brahman represents the ultimate reality in all existence, wherein each individual's consciousness materializes from a unitary consciousness suffused throughout the universe.


268. This statement is not true in general; it is true in so far as the human brain is healthy. The brain is enclosed in a closed vault (i.e., the skull). If the brain becomes edematous (i.e., excessive accumulation of serous fluid) due to, for example, a traumatic brain injury, then this will increase the pressure (the intracranial pressure) inside this closed vault. If the intracranial pressure becomes too large, then the brain will be compressed and this can result in serious injury if not death. In cases of intracranial pathology (e.g., brain tumors, traumatic injury to the brain, and bleeding in the brain) there will be increased edema, and hence, increased intracranial pressure as well. This is exacerbated by increased carbon dioxide as this increases blood flow to the brain and increases the edema fluid load.

269. When patients lose consciousness other parts of the brain are still functional (heart rate control, ventilation, oxygenation, etc.), and hence, the development of biological neural network models that exhibit partial synchronization is critical. In particular, models that can handle synchronization of subsets of the brain with the non-synchronized parts firing at normal levels is essential in capturing biophysical behavior. For further details see [4,5].


272. All living things are chronognostic as they adopt their behavior to a dynamic (i.e., changing) environment. Thus, behavior has a temporal component, wherein all living organisms perceive their future relative to the present. Behavioral anticipation is vital for survival, and hence, chronognosis is zoocentric and not only anthropocentric.


274. As in classical mechanics, which serves as an approximation theory to the theory of relativity, classical electromagnetism predicated on Maxwell’s field equations is an approximation to the relativistic field theory of electrodynamics describing how light and matter interact through the exchange of photons.


Even though Newton stated that “A change in motion is proportional to the motive force impressed and takes place along the straight line in which the force is impressed”, it was Euler, almost a century later, that expressed this statement as a mathematical equation involving force and change in momentum. Contrary to mainstream perception, the mathematical formalism and rigor used by Newton was elementary as compared to the mathematics of Euclid, Apollonios, and Archimedes. One needs to only compare how Newton presents his formulation on the limit of the ratio of two infinitesimals, which he calls the “ultimate proportion of evanescent quantities”, in [166] (Book I, Section I), with Archimedes’ Proposition 5 On Spirals [280] (pp. 17–18), where he uses infinitesimals of different orders to determine the tangential direction of an arbitrary point of a spiral. This comparison clearly shows that Newton lacked the mathematical sophistication developed two thousand years earlier by Hellenistic mathematicians. This is further substantiated by comparing the mathematical formalism in [166] with [280–283].

Later Planck and Einstein modified this view of the nature of light to one involving a wave-particle duality. Einstein’s photon (particle) theory of light [285] asserted that energy flow is not continuous but rather evolves in indivisible packets or quanta, and light behaves at times as a wave and at other times as a particle. And this behavior depends on what an observer chooses to measure. This wave-particle duality of the nature of light led to the foundations of quantum physics, the Heisenberg uncertainty principle, and the demise of determinism in the microcosm of science.

The most famous of these experiments was the Michelson-Morley experiment performed in the Spring and Summer of 1887 [286]. In an attempt to detect the relative motion of matter through the stationary luminiferous aether and find a state of absolute rest for electromagnetic phenomena, Michelson and Morley compared the speed of light in perpendicular directions. They could not find any difference in the speed of electromagnetic waves in any direction in the presumed aether. Over the years, many Michelson-Morley type experiments have been performed with increased sensitivity and all resulting in negative results ruling out the existence of a stationary aether.


The Lorentz transformations only describe transformations wherein the spacetime event at the origin of the coordinate system is fixed, and hence, they are a special case of the Poincaré group of symmetry transformations which include translation of the origin.

297. An affine connection is a geometric object (e.g., points, vectors, arcs, functions, curves) defined on a smooth manifold connecting tangent spaces that allows differentiability of tangent vector fields. The Levi-Civita (affine) connection is a torsion-free connection on the tangent bundle of a manifold that preserves a semi-Riemannian metric.

298. Since a zero gravitational field is a well-defined field which can be measured and changed, it provides a gravitational field (i.e., the zero gravitational field) to which acceleration can be relative to. Thus, special relativity can be viewed as a special case of general relativity for which the gravitational field is zero.


302. Even though Einstein and Planck are credited with developing the first relativistic thermodynamic theory, it was von Mosengeil [303] who was the first to arrive at the relativistic temperature transformation expression of (11).


337. Since in relativity theory two separate events that are simultaneous with respect to one reference frame are not necessarily simultaneous with respect to another reference frame, length measurements as well as time interval measurements depend on the reference frame of the observer, and hence, are relative. Thus, location and time measurement devices (clocks) need to be compared, calibrated, and synchronized against one another so that observers moving relative to each other measure the same speed of light. For further details see [291].
338. A notable exception is [323, 333] which will be discussed in Section 15.
339. Planck, M. *Vorlesungen Über Die Theorie der Wärmestrahlung*; J. A. Barth: Leipzig, Germany, 1913. (In German)
341. If one were to adopt Einstein’s famous maxim in that a theory should be “as simple as possible, but no simpler”, then one would surmise that temperature is a relativistic invariant.
342. For sufficiently light molecules at high temperatures (74) becomes $E_{KE} = 3NkT$ [343].

© 2017 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).