# Where is the entropy challenge?

George N. Hatsopoulos[*] and Gian Paolo Beretta[†]

[*]*MIT, Cambridge, MA 02139*
[†]*Università di Brescia, Italy*

**Abstract.** We discuss past, present, and future entropy challenges due to the fundamental questions that arise when the Second Law of Thermodynamics is confronted with Classical and Quantum Mechanics.

## FOREWORD AND SUMMARY

Thermodynamics is one of the few sciences involving fundamental concepts that are controversial and have been so for well over one century. The basic issues are the range of validity of the Second Law of Thermodynamics, the meaning of entropy, and the origin of irreversibility. Over the years, several schools of thought emerged each of which prevail among sizable groups of scientists and engineers, even today. Some of these points of view are as follows:

1. The Second Law of Thermodynamics and the concept of entropy are approximations applicable to macroscopic systems that appear to be in equilibrium because the observer lacks the ability to detect the actual microscopic state of the system.[1]

2. The Second Law of Thermodynamics and the concept of entropy apply rigorously to all systems, whether those that appear to be in equilibrium or not, but only to the system-observer combination.[2] It follows that entropy is not an intrinsic property of systems; rather it is a property of the system-observer combination.

3. The only quantal dispersions that are inherent to the state of any system are those particular dispersions described by a "pure" wave function, i.e., an idempotent density operator. Moreover, the universe as a whole is in a pure state.[3] The Second Law limitations we observe in some system, e.g. system A, are due to entanglement of A with some other system B such that A plus B is in a pure state while each of A and B, when viewed locally, are in mixed states.

4. Any system can assume states that possess inherent quantal dispersions including those described by a "pure" wave function, i.e., those described by an idempotent density operator, and those described by a non-idempotent density operator. As

---

[1] The "classical" view.

[2] The "information theory" view.

[3] The "quantum information" view, discussed also in the papers by Bennett and Lloyd in this volume.

a result, the Second Law of Thermodynamics and the concept of entropy apply rigorously to all systems.[4]

In this paper, we discuss the many fundamental questions that in our view still challenge many details of each of these points of view.

# ENTROPY CHALLENGES IN THE 19th CENTURY

Entropy is a scientific concept introduced by R.J.E. Clausius in 1850 [1] to reconcile the conclusions reached by N.L.S. Carnot in 1924 [2] concerning the motive power of heat with those reached by J.P. Joule in 1848 [3] concerning the equivalence of heat and work.

The motivation behind Carnot's scientific effort was to find the basis of improving Watt's steam engine, invented 60 years earlier. Unlike steam engines in the past, Watt's engine was the first steam engine that did not consume water—it only received heat and produced work. Thus, it was the first true "*Heat Engine*". At that time the scientific community thought that heat was a fluid called *caloric* and that Watt's engine was nothing but a turbine that takes that fluid from a high level (a boiler at high temperature), produces useful work, and ejects it at a lower level (a condenser at lower temperature).

Carnot devised a reversible engine operating in a different cycle[5] than Watt's engine. The ratio of the net work output to the heat input, called the efficiency of the cyclic engine, is proportional to the difference between the temperatures of the heat source and the heat sink. Carnot asserted that it is the largest such ratio that any engine operating between these two temperatures can ever reach.[6] This assertion is known as the Carnot principle. It also implies that (Kelvin-Planck statement of the Second Law): a cyclic engine that produces work (and no other effects) by exchanging heat with a single reservoir is impossible. Such an engine is called a *perpetual motion machine of the second kind* (PMM2).[7]

During the period from 1840 to 1848, James Prescott Joule showed experimentally that heat and work could produce the same effect on bodies when used in a fixed proportion. Thus, in a cyclic process, such as that of a cyclic engine, the net work produced must be proportional to the net heat received. He concluded [3] that either heat or work results in a change of something 'stored' in the bodies, which is conserved.

---

[4] The "Hatsopoulos, Gyftopoulos, and Keenan" view.

[5] In the Carnot cycle the engine's working substance undergoes four successive changes: It receives heat (from the heat source) while expanding at high temperature, delivers work during a reversible adiabatic (no heat) expansion, rejects heat (to the heat sink) during a compression at low temperature, and finally receives work during a reversible adiabatic compression.

[6] It is noteworthy that the essential new element in Carnot's reasoning, which grants the maximal efficiency for the given source and sink temperatures, is not just the particular cycle he devised for the working substance, but the hypothesis of reversibility he conceived for his engine. Any other cyclic engine operating reversibly would achieve the same efficiency, regardless of the cycle undergone by the working substance.

[7] This definition is analogous to that of a *perpetual motion machine of the first kind* (PMM1) which instead would produce more work than the net energy it receives form sources and sinks.

We now call that something *energy*.

In 1849 Lord Kelvin [4], a Scottish engineer, pointed out the conflict between the caloric basis of Carnot's argument in which heat (caloric) is conserved and the conclusion reached by Joule in which the sum of work and heat is conserved. Moreover, Joule's theory poses no limits on how much of the heat can be transformed into work, whereas Carnot's theory does. One year later, in 1850, Clausius [1] reconciled Carnot's principle with Joule's result by introducing the concept that bodies possess a property, which he finally called *entropy* in 1865 [5], having the following characteristics: In the absence of heat interactions with other bodies, it either remains constant if the body undergoes a reversible process, or increases. During heat interactions, on the other hand, entropy is transferred to or from a body in proportion to the heat transferred. It is this later characteristic that limits the efficiency of any work-producing cyclical engine, as required by Carnot's principle, from which Clausius concluded that (Clausius statement of the Second Law): *without expenditure of work (or some equivalent external effect), a transfer of energy from a body at any temperature to a body at higher temperature is impossible*.

Joule's finding directly relates the First Law of Thermodynamics (and the resulting principle of energy conservation) to Newton's laws of motion. On the other hand, relating the Second Law to Newtonian Mechanics proved more difficult. Although Statistical Mechanics relates the thermodynamic entropy to Mechanics, a major conflict between the two sciences remains. It was pointed out by Maxwell and is illustrated very clearly by what has come to be known as Maxwell's demon [6].

This conflict results from the fact that although Mechanics allows under all circumstances the extraction of any fraction of the energy of any physical system confined within a given volume in the form of work, in some states the Second Law limits that fraction, depending on the value of the entropy (a property 'possessed' by all systems in any specified condition). Only if the entropy of a system has the lowest value possible at the given energy, can one extract all its energy in the form of work. Under that condition, the laws of Mechanics and Thermodynamics become identical.

Until the mid twentieth century, the scientific community believed that the Second Law is an approximation relating only to macroscopic systems which observers perceive to be in equilibrium, but which in reality are continuously going through a very large number of microstates following Newton's laws of motion. According to this view, the properties of the macroscopic system we observe are only time averages. For them entropy relates to the probability of finding the system over time in any one particular microstate. It follows that entropy is only meaningful for a macroscopic system we perceive to be in equilibrium not because it actually is in equilibrium but because we lack the ability to see changes over time. As to the possibility of extracting the energy of such a system in the form of work, we can do so only if we get the ability to observe the molecular details of the system as Maxwell's demon could. In *The Dynamical Theory of Gases*, Sir James Jeans writes [7]: "Thus Maxwell's sorting demon could do in a very short time what would probably take a very long time to come about if left to the play of chance. There would, however, be nothing contrary to natural laws in the one case any more than in the other". The implied bottom line is that a PMM2 is feasible although beyond the technology at that time.

# ENTROPY CHALLENGES IN THE 20th CENTURY

By the end of the last century most scientists have been brought to believe that the conflict between Thermodynamics and Mechanics was resolved by Szilard in his famous paper of 1929 [8], and Brillouin who in 1956 [9] combined Szilard's concept with the information theory developed by Shannon in 1948 [10].

Szilard's premise may be summarized as follows: We shall accept the proposition that it is possible to construct mechanical devices that make use of any one fluctuation of a system in stable equilibrium to produce work. Moreover, we shall accept the Second Law in the form that no net positive work may be obtained on the average from a system in stable equilibrium without producing any other net average effects on the environment. From these assumptions, we conclude that any instrument (or demon) used to identify any given fluctuation of a system in stable equilibrium will absorb a quantity of work which is at least as much as the work that may be obtained from the fluctuation. More recently, based on Landauer's work [11], Bennett [12] pointed out that if Szilard's engine consists of a box with a single particle of gas, the demon puts the piston in, measures where the particle is and gains 1 bit of information. He, then, uses it to decide how to move the piston so that by expansion, he draws at most $k_B T \ln 2$ of work, but the price is that at least $k_B T \ln 2$ of work must be spent to erase the 1 bit from his memory and get back to its initial state.

In his *Treatise on Thermodynamics* (1927) M. Planck states the Second Law of Thermodynamics as follows [13]: "*It is impossible to construct an engine which will work in a complete cycle, and produce no effect except the raising of a weight and the cooling of a heat reservoir.*" If we assume that what he means by "*heat reservoir*" is a system in stable equilibrium sufficiently large (or passing through two- or three-phase states) so that changes in its energy do not affect its temperature and moreover that such systems do exist, then the following statement is a corollary of the Second Law:[8]

*There exist states for any system such that the largest amount of energy that can be extracted from it in the form of work without any other changes in the environment is less than 100% of the system's energy (above the ground energy for the given values of amounts of constituents and parameters). Such a fraction of the system's energy is an intrinsic property of the given state of the system.*[9]

For a system in combination with a thermal reservoir, this property is called "*availability*".[10] The difference between the energy $E$ of a system and its availability $\Omega^R$, times a constant $C_R$ solely related to the reservoir $R$ with respect to which availability is measured, defines another intrinsic property called "*entropy*" which can be proved to be independent of the reservoir used for its definition,[11] and valid also for nonequilibrium

---

[8] A full account of this point is cited in Refs. [14] and [15].

[9] In Refs. [16, p.130] and [15] this property is called "adiabatic availability", and is denoted by the symbol $\Psi$.

[10] More precisely, in Ref. [16, p.132] this property is called "available energy with respect to a reservoir", denoted by the symbol $\Omega$. In Ref. [15] it is denoted by $\Omega^R$ to expose its dependence on the reservoir.

[11] A detailed proof of this independence is given in Ref. [15, par.7.4, p.108].

states.[12]

On the other hand, the Szilard proposition that the Second Law relates to the ignorance that an observer has about the actual intrinsic state of a system, and therefore that entropy is not an intrinsic property of the system can be proven wrong by means of the following experiment:

*Alice places a dozen identical batteries in a large room at some constant temperature. She charges some and totally discharges the others. The charged ones have finite availability and those discharged zero availability. Then Bob enters the room. He knows nothing about which batteries are charged. Bob can easily identify the charged ones and determine their availability by slowly discharging each of them and measuring the heat transferred to the room during the discharge operation. From these two parameters and the room temperature the observer can calculate the change of entropy during the discharging process. By repeating the experiment several times and getting the same answer, a result that any engineer will predict, Bob, with no need of receiving any information from Alice, will have separated the charged from the discharged batteries. Availability, and hence entropy, is an intrinsic objective property of the batteries.*

In fact, there are numerous examples that illustrate that the subjective informational interpretation of the Second Law makes no physical sense. For example, consider a box separated into two halves by a partition. One half contains a gas in equilibrium at some temperature, the other is a vacuum.[13] The entropy of the box can be easily calculated. If we lift the partition and wait long enough until equilibrium is reached, the energy of the gas will be the same as before but its entropy will be much larger since there are more microstates available to it. Now let us ask what the entropy of the system is at in-between times. Some will answer that since at those times the system is not in equilibrium, entropy is indefinable. On the other hand, it is common sense to say that more work can be extracted from the system at those times, than when the system reaches equilibrium, i.e., its availability at those times is greater. The change of availability upon removal of the partition is an objective characteristic of the state of the system and has nothing to do with the observer. Removing the partition makes more microstates accessible and increases the availability instantly. The entropy instead is not affected instantly.[14] Only when the gas begins to diffuse into the newly available volume, the entropy increases and the availability decreases, until the gas reaches a new stable equilibrium, the entropy achieves the maximal value compatible with the new volume (and the initial energy and amount of particles), and the availability achieves a minimum (the adiabatic availability falls to zero). Innumerable other phenomena similar to this example occur in Nature whether we observe them or not, whether we reproduce them under controlled conditions in a lab or they happen spontaneously and unobserved. They are part of the "objective" empirical world—the "physical reality" [18]—that our physical theories set out to regularize. The observer plays no role, and there is nothing subjective.

---

[12] See Refs. [15], [16], [17], and the paper by Zanchini and Beretta in the present volume.

[13] On trying to define entropy for non-equilibrium states, see the last part of the interview with Joseph Keenan on May 13, 1977 published as his Autobiographical Notes in the present volume.

[14] Of course, here we mean the nonequilibrium entropy, as we just defined it.[12]

The advent of the wave theory of matter (Quantum Mechanics) and, specifically, the introduction in 1927 of Heisenberg's principle of indeterminacy [19] raised great hopes that the paradox posed by Maxwell's demon might be resolved and, moreover, that a complete proof of the Second Law of Thermodynamics could be obtained based only on quantum-mechanical principles. Slater [20] attempted the former and Watanabe [21] the latter. Both attempts failed. Demers [22] proved that dispersions associated with "pure" wave functions[15] are insufficient to account for the implications of the Second Law.

The second law requirements, however, need not conflict with the foundations of Quantum Mechanics. Unlike Newtonian Mechanics, Quantum Mechanics stipulates that the state of a physical body at a given time can be described, at best, by a cloud of probabilities to find its particles with specific positions and momenta. In other words, quantum theory postulates that the *state* of any physical system incorporates *irreducible quantal dispersions* that are inherent to it. As shown by Demers, the conflict observed by Slater arises only if we assert that the **only** states that a system can assume are those having quantal dispersions describable by an idempotent density operator, and therefore a "pure" wave function.[16] If, on the other hand we assert the possibility that systems can assume states having quantal dispersions describable by a non-idempotent density operator, the Second Law becomes consistent with Quantum Theory. Neither the Second Law nor either of these two assertions could be proved right, and none has been proved wrong.

It is conceivable that the Second Law is proved wrong. All it will take is a single experiment that violates it. For example, if we find a way to continuously produce work at the only expense of extracting energy from a single large environment in stable equilibrium. The same conclusion applies to the First Law (and the principle of conservation of energy that it entails). In both cases, however, the only reason we surmise their validity is that they have never been violated despite continuous efforts[17] resulting from the enormous societal and financial benefits that the world would derive from their violation.

On the other hand, the assertion that **all** the states that any physical system can assume

---

[15] Technically, we call a density operator "idempotent" or "pure" if it can be written as $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle$ is some unit vector in the Hilbert space $\mathscr{H}$ of the system, the quantum mechanical "pure state" vector. Clearly, $\rho^2 = \rho$. If we denote by $Q$ the operator associated with measuring the positions of every particle in the system and by $Q|\boldsymbol{q}\rangle = \boldsymbol{q}|\boldsymbol{q}\rangle$ its eigenvalue problem, the *position-representation "pure" wave function* is $\psi_Q(\boldsymbol{q}) = \langle\psi|\boldsymbol{q}\rangle$ and the probability to find the particles with positions between $\boldsymbol{q}$ and $\boldsymbol{q} + d\boldsymbol{q}$ is given by $\psi_Q(\boldsymbol{q})\psi_Q^*(\boldsymbol{q})\,d\boldsymbol{q}$. Similarly, if we denote by $P$ the operator associated with measuring the momenta of every particle in the system and by $P|\boldsymbol{p}\rangle = \boldsymbol{p}|\boldsymbol{p}\rangle$ its eigenvalue problem, the *momentum-representation "pure" wave function* is $\psi_P(\boldsymbol{p}) = \langle\psi|\boldsymbol{p}\rangle$, and the probability to find the particles with momenta between $\boldsymbol{p}$ and $\boldsymbol{p} + d\boldsymbol{p}$ is $\psi_P(\boldsymbol{p})\psi_P^*(\boldsymbol{p})\,d\boldsymbol{p}$. If instead $\rho^2 \neq \rho$, then we call the density operator "non-idempotent" or "mixed", and the corresponding position-representation and momentum-representation *"mixed" wave functions* are respectively given by $\psi_Q(\boldsymbol{q}) = \langle\boldsymbol{q}|\rho|\boldsymbol{q}\rangle$ and $\psi_P(\boldsymbol{p}) = \langle\boldsymbol{p}|\rho|\boldsymbol{p}\rangle$.

[16] More precisely, by an idempotent density operator, as defined above and further discussed in the next sections.

[17] For many years during Professor Keenan's career he reviewed dozens of requests each year to find the error in proposed inventions that violated the Second Law. Some of these requests came from the US Patent Office. Such requests stopped when the Patent Office adopted the policy of rejecting all patent applications for inventions clearly violating the Second Law.

are describable by a pure wave function is pure speculation because it is impossible to prove it wrong by means of any finite number of experiments. All we know is that the states of systems for which we have experimentally identified quantum dispersions are describable by a pure wave function. For these states the Second Law is irrelevant.

The Kelvin-Planck, the Clausius, and the Carathéodory statements of the Second Law, can all be shown[18] to be logical consequences of the following more general statement of the Second Law proposed by Hatsopoulos and Keenan in 1965 [14] (see also [15, 16]):

*Among all the states of a system[19] that have a given value E of the energy and are compatible with a given set of values **n** of the amounts of constituents and **β** of the parameters of the external forces (i.e., parameters of the Hamiltonian), there exists one and only one stable equilibrium state.[20]*

This statement of the Second Law brings out very clearly the apparent conflict between Mechanics and Thermodynamics, a contrast that for over a century has been perceived as paradoxical. In fact, within Mechanics, classical or quantum, the following so-called *minimum energy principle* applies: *Among all the states of a system that are compatible with a given set of values **n** of the amounts of constituents and **β** of the parameters of the external forces (or the Hamiltonian), there exists one and only one stable equilibrium state, that of minimal energy.* Comparing the two assertions, if we insist that the two theories of Nature contemplate the same set of states, than a paradox arises, because, for given **n** and **β**, Mechanics asserts that there is only one stable equilibrium state, whereas Thermodynamics that there is one for each of the infinite values energy can take.

The paradox is removed if we admit that the "pure" states contemplated by Quantum Mechanics are only a subset of those contemplated by Thermodynamics. This resolving assumption was very controversial when Hatsopoulos and Gyftopoulos in Ref. [16] first introduced it. However today—more as a byproduct of the more recent vast literature on quantum entanglement and quantum information than as a result of thermodynamic reasoning—an assumption to this effect has been included in the postulates of quantum

---

[18] Explicit proofs of this assertion can be found in Ref. [15, p.64-65 (Kelvin-Planck), p.133-136 (Clausius), and p.121 (Carathéodory)].

[19] The usage of the term *system* implies the following restriction: we may call system any collection of elementary constituents provided they are not subjected to forces that depend explicitly on coordinates of constituents outside the collection. To emphasize this important restriction, without which for example it would be impossible to define energy as an additive property, Ref. [16, p.18] defined precisely and adopted the term *separable system*. Here and in Ref. [15] the term *separable* in this sense is not adopted, to avoid confusion with the concept of separability related to absence of quantum entanglement, which became widespread in more recent years. However, within Thermodynamics as well as Classical or Quantum Mechanics, any well-defined model of a physical system is in terms of external forces and an Hamiltonian that depends only on geometrical or other kind of parameters, and on degrees of freedom internal to the system.

[20] A *stable equilibrium state* is an equilibrium state that cannot be altered to a different state without leaving net effects in the environment. It can always be restored if the state of the environment undergoes a cycle [16], or, equivalently, it can be changed only by interactions that either leave net effects in the environment of the system, or change the values of amounts of constituents and parameters to an incompatible set of values [15].

theory,[21] and the discussion of its relations with Thermodynamics is flourishing.[22]

## ENTROPY CHALLENGES FOR THE 21st CENTURY: THE SECOND LAW AND QUANTUM STATES

The fundamental change that Quantum Theory introduced in Mechanics is the conclusion that, ***however accurate the measuring instruments may be, there exists no state of any physical system without at least one observable for which the outcome of a single act of measurement cannot be predicted with certainty, and all that can be predicted are the probabilities of occurrence of a spectrum of possible outcomes***. Thus, we can say that "*the Uncertainty Principle is an intrinsic characteristic of matter*". Since this assertion is contrary to Newtonian Mechanics, we conclude that the latter is an approximation (very good for macroscopic systems).

To determine empirically the probability distribution for *one* observable, we must repeat the measurement procedure that defines that particular observable, on an *ensemble* of identical replicas of the system, all identically prepared, and collect the statistics of the outcomes. Von Neumann (1932) [26] recognized that the resulting probability distribution, which is a feature of the *preparation* used to generate the ensemble, may not be considered an intrinsic feature of each individual member of the ensemble, unless the preparation, and hence the ensemble, is *homogeneous*.[23] Indeed, only when there is no conceivable decomposition of the ensemble into different subensembles, the statistics of the ensemble cannot originate but from the individual state of each and every one of its members.

To determine empirically the *state* of a system at one time, we must measure, on a homogeneous ensemble of identically prepared identical replicas of the system, the probability distributions for a complete set (quorum) of independent observables,[24] sufficient to determine the probability distributions of all other conceivable observables.

Paul A.M. Dirac [28] established the modern formalism of Quantum Theory in 1930. He unified the two major theories of atomic phenomena [29]: matrix mechanics, developed by Werner Heisenberg with help of Max Born and Pascual Jordan around late 1925-early 1926, and wave mechanics developed by Erwin Schrödinger in early 1926 based on ideas set forth by Louis de Broglie in 1923. Since then, however, the master-

---

[21] Compare for example the postulates of quantum theory as stated, e.g., in the recent Ref. [23] with those stated by Park and Margenau in 1968 [24].

[22] See for example Ref. [25] and references therein (which unfortunately do not include [16]).

[23] An ensemble is homogeneous if and only if no partitioning scheme can possibly subdivide it into subensembles that would yield measurement statistics different from the original ensemble. An example of a heterogeneous ensemble is one resulting from the random statistical composition of two *different* preparations $\Pi_1$ and $\Pi_2$ (different meaning that they yield different statistics for at least one observable): for example, we toss a coin and, depending on whether we get head or tail we prepare the system according to $\Pi_1$ or $\Pi_2$, respectively.

[24] The *empirical determination of a quantum state* has been recently called **quantum tomography**. In a seminal series of papers, Park and Band (1970-1971) [27] devised elegant systematic rules to construct a **quorum** of observables whose mean values are sufficient to fully determine the probability distributions of all other observables and, hence, the state of the system.

ing of the formalism, its conceptual implications about the nature of physical reality, its experimental validation, and its applications has been continuously evolving. As put in a recent paper [30]: "It is astonishing that after over 60 years of study, the quantum formalism has only recently revealed us new possibilities due to entanglement processing being a root of such new quantum phenomena as quantum cryptography with the Bell theorem, quantum dense coding, quantum teleportation, quantum computation, etc. It shows how important it is to recognize not only the structure of the formalism itself, but also the potential possibilities encoded within. In spite of many wonderful experimental and theoretical results on entanglement, there are still difficulties in understanding its many faces." It is conceivable that entanglement may also be the fundamental reason for the existence of the irreducible quantal dispersions that in Ref. [16] were first shown to be all that is needed to reconcile Mechanics and Thermodynamics, independently of the question of irreversibility.

Initially von Neumann [26] introduced density operators to represent statistics of measurement results obtained from *heterogeneous* ensembles, as a logical way to extend the Statistical-Mechanics notion of distribution over a set of possible "microstates" to the quantum domain [31]. If the measurement statistics from preparations $\Pi_1$ and $\Pi_2$ are represented by density operators $\rho_1$ and $\rho_2$, respectively, their random statistical composition with weights $w$ and $1-w$ is represented by density operator $\rho = w\rho_1 + (1-w)\rho_2$. Thus, density operators form a convex set whose extreme elements are the pure states, the only ones that cannot be decomposed as a weighted sum of two *different* density operators, and hence the only ones that may represent *homogeneous* ensembles. For a two-level system, pure states map to a three-dimensional unit sphere (the Bloch sphere), and mixed density operators to the enclosed unit ball. In general, pure states map one-to-one with the unit norm vectors $|\psi\rangle$ in the Hilbert space of the system, the corresponding density operators being $\rho = \rho^2 = |\psi\rangle\langle\psi|$ (i.e., the idempotent, one dimensional projectors onto the linear span of the state vector $|\psi\rangle$).

The essential new features of Quantum Theory versus Classical Mechanics can be brought out with no loss of generality by focusing on physical observables that have a binary spectrum, i.e., that upon measurement yield only two possible outcomes, such as 'one' or 'zero' (or 'yes' or 'no'). In fact, all other observables can be expressed as combinations of binary observables. In Classical Mechanics, a binary observable is represented by a unit valued function $P$ defined over the system's phase space $\Omega$ which induces a disjoint partition $\Omega = \Omega_P \bigcup \Omega_{1-P}$ into a region $\Omega_P$ where the measurement outcome is always 'one' and the region $\Omega_{1-P}$ where the outcome is always 'zero'. Since the 'answers' are either always 'one' or always 'zero', the intersection $\Omega_P \bigcap \Omega_{1-P}$ is empty by definition. In Quantum Theory, instead, a binary observable is represented by a projection operator $P$ onto some subspace $\mathscr{H}_P$ of the system's Hilbert space $\mathscr{H} = \mathscr{H}_P \bigoplus \mathscr{H}_{1-P}$, so that for state vectors $|\psi\rangle$ that belong to subspace $\mathscr{H}_P$ (where $P|\psi\rangle = |\psi\rangle$) the measurement outcome is always 'one', for state vectors in the orthogonal complement subspace $\mathscr{H}_{1-P}$ (where $P|\psi\rangle = 0$ or equivalently $(I-P)|\psi\rangle = |\psi\rangle$) the outcome is always 'zero', but in addition there are state vectors that have nonzero orthogonal components both in $\mathscr{H}_P$ and $\mathscr{H}_{1-P}$ (i.e., $|\psi\rangle = P|\psi\rangle + (I-P)|\psi\rangle$, with $0 < \langle\psi|P|\psi\rangle < 1$). For the latter state vectors, the outcome is not predictable, but the probability of getting 'one' is well defined and given by $p = \langle\psi|P|\psi\rangle$ or, equivalently, $\text{Tr}(P|\psi\rangle\langle\psi|)$. For a statistical mixture of two homogeneous ensembles, respectively

represented by state vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ (not necessarily orthogonal to each other), the von Neumann recipe assigns the density operator $\rho = w|\psi_1\rangle\langle\psi_1| + (1-w)|\psi_2\rangle\langle\psi_2|$, and the probability of getting 'one' is given by $p = \text{Tr}(P\rho) = w\,p_1 + (1-w)\,p_2$, as it should be.

The von Neumann construction has been extremely successful because it prompted the "translation" of the Gibbs-Boltzmann canonical distribution of equilibrium Statistical Mechanics into the corresponding density operators $\rho = \exp(-H/k_{\text{B}}T)/\text{Tr}[\exp(-H/k_{\text{B}}T)]$ where $H$ is the system's Hamiltonian operator; and the subsequent restriction of this to the symmetric and antisymmetric subspaces of the Hilbert space of an assembly of indistinguishable particles succeeded in regularizing Bose-Einstein and Fermi-Dirac statistics, respectively.

However, already in 1936 Schrödinger [32, 33] expressed his discomfort with von Neumann's conceptual construction, arising from the fact that the same mixed density operator can be resolved in different statistical mixtures of incompatible sets of component pure states. This means that the same mixed density operator $\rho$ represents an infinite number of statistical mixtures of *different* pairs of pure states,[25] i.e., there are infinite pairs of different state vectors $\{|\psi_1\rangle, |\psi_2\rangle\}$, $\{|\psi_3\rangle, |\psi_4\rangle\}$, etc. such that $\rho = w_{12}|\psi_1\rangle\langle\psi_1| + (1-w_{12})|\psi_2\rangle\langle\psi_2| = w_{34}|\psi_3\rangle\langle\psi_3| + (1-w_{34})|\psi_4\rangle\langle\psi_4| = \ldots$. This leads to the following ***paradox***: according to the von Neumann construction, the first decomposition implies that the (heterogeneous) ensemble is $\{\{w_{12}, 1 - w_{12}\}, \{|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|\}\}$ and therefore *the individual state of the systems in the ensemble can only be either $|\psi_1\rangle$ or $|\psi_2\rangle$*, but the second decomposition implies the ensemble $\{\{w_{34}, 1 - w_{34}\}, \{|\psi_3\rangle\langle\psi_3|, |\psi_4\rangle\langle\psi_4|\}\}$ and therefore that *the individual state of the systems in the ensemble can only be either $|\psi_3\rangle$ or $|\psi_4\rangle$*, which clearly contradicts the preceding conclusion. So, if the language of individual states is applied to these ensembles, as Park put it in 1968 [34]: "Immediately a quantum monster is born: a single system concurrently 'in' two states $|\psi_1\rangle$ *and* $|\psi_3\rangle$. ... Thus the concept of individual quantum state is fraught with ambiguity and should therefore be avoided in serious philosophic inquiries concerning the nature of quantum theory."

As argued in Ref. [33], the fundamental assumption introduced by Hatsopoulos and Gyftopoulos in Ref. [16] in order to remove the conflict between Mechanics and Thermodynamics (one stable equilibrium state versus many, for given $\boldsymbol{n}$ and $\boldsymbol{\beta}$), removes also the Schrödinger-Park paradox about individual states. All it takes is to accept that in general the density operators be associated with the homogeneous ensembles (or preparations)—not the heterogeneous ones. Each density operator is raised to represent a 'true' individual state, true in the same sense as the state vectors represent individual states within standard von Neumann Quantum Mechanics.

This key assumption resolves the paradox and at the same time unifies Mechanics and Thermodynamics into a single general uncontradictory theory. However, it is also a source of foundational questions.

---

[25] This property has a simple geometrical proof in the Bloch ball/sphere representation, for a two level system. Through every point in the interior of the unit sphere (which represents a mixed density operator) there pass an infinity of chords; each chord identifies a possible resolution of the density operator as a mixture of the pure states represented by the end points of the chord.

The assumption that there are preparations which produce homogeneous ensembles of *independent* and separable systems whose measurement statistics require mixed density operators, was contradictory with the standard postulates [24] of Quantum Theory at the time it was introduced, and it still is, if the condition of independence is maintained. It is true that Quantum Theory is currently formulated [23] by postulating that homogeneous ensembles of separable systems are represented by pure as well as mixed density operators. But the current general understanding is that mixed density operators are required for homogeneous ensembles only when every system $A$ in the ensemble is entangled with some external system $B$. This means that the $A$-systems ensemble can be purified, because there exists a correlated $B$-systems ensemble somewhere in the Universe, such that the homogeneous ensemble of composite systems $AB$ is represented by a pure density operator $|\psi_{AB}\rangle\langle\psi_{AB}|$. Mathematically, such existence is a granted by the Schmidt decomposition theorem together with the fact that for any local binary observable $P_A \otimes I_B$ of system $A$, the outcome probability of the 'one' answer is $p_A = \langle\psi_{AB}|P_A\otimes I_B|\psi_{AB}\rangle = \text{Tr}(\rho_A P_A)$ where $\rho_A = \text{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|)$, therefore the reduced density operator $\rho_A$ 'contains' all local probability distributions.[26]

The following question arises. There seem to be many ways to prepare an ensemble described by a given mixed density operator $\rho_A$, many ways that appear physically very different from one another:

1. system $A$ is an entangled subsystem of a member system $AB$ of a "homogeneous" ensemble described by the idempotent density operator $|\psi_{AB}\rangle\langle\psi_{AB}|$ with $\text{Tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|) = \rho_A$; system $B$ may be anywhere, as $A$ need not be interacting with $B$; but they did interact in the past, when the quantum correlation (*entanglement*) has been established;

2. system $A$ is a member of a *heterogeneous* ensemble $\{\{w_{12}, 1-w_{12}\}, \{|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|\}\}$ with $w_{12}|\psi_1\rangle\langle\psi_1| + (1-w_{12})|\psi_2\rangle\langle\psi_2| = \rho_A$, $0 < w_{12} < 1$, and $\langle\psi_1|\psi_2\rangle = 0$, i.e., a statistical mixture of homogeneous components represented by

---

[26] For the overall system $AB$, the outcome 'one' of a binary observable $P_A \otimes P_B$ implies that both local measurements on $A$ and $B$ gave "one", the (joint) probability being $p_{AB} = \langle\psi_{AB}|P_A\otimes P_B|\psi_{AB}\rangle$. If $A$ and $B$ are entangled, $p_{AB}$ differs from $p_A p_B$, and this happens whenever there is no pair of vectors $|\alpha\rangle$ in $\mathscr{H}_A$ and $|\beta\rangle$ in $\mathscr{H}_B$ such that $|\psi_{AB}\rangle = |\alpha\rangle \otimes |\beta\rangle$. The entanglement is maximal when $p_{AB} = p_A = p_B$, i.e., when every time $A$ measures 'one' also $B$ measures 'one'. This striking, purely quantum effect has generated much debate over the years, mainly because it defeats classical reasoning. Indeed, if we are sure that the state of the pair $AB$ is maximally entangled, then upon obtaining 'one' from $A$ we are *certain* that also $B$ would yield 'one', even if $A$ and $B$ are far apart and do not communicate. What is striking is that, before the measurement, it is impossible to predict what will be the outcomes, so according to our classical thinking it appears that either

1. $A$ and $B$ instantly communicate with each other in order to *agree on* which of the two possible answers to yield (thus violating the impossibility of instantaneous or faster-than-light-speed communication), or else

2. there is some additional (hidden variable) local feature, additional to the state vector $|\psi_{AB}\rangle$, that each subsystem consults when subjected to a measurement, where it is written a priori which answer they must give (thus violating the principle that Heisenberg indeterminacy is irreducible and intrinsic in the nature of every quantum state).

*orthogonal* state vectors;

3. system A is a member of a *heterogeneous* ensemble $\{\{w_{34}, 1 - w_{34}\}, \{|\psi_3\rangle\langle\psi_3|,$ $|\psi_4\rangle\langle\psi_4|\}\}$ with $w_{34}|\psi_3\rangle\langle\psi_3| + (1 - w_{34})|\psi_4\rangle\langle\psi_4| = \rho_A,\ 0 < w_{34} < 1,$ and $\langle\psi_3|\psi_4\rangle \neq 0$, i.e., a statistical mixture of homogeneous components represented by *non-orthogonal* state vectors;

4. system A is a member of a *homogeneous* ensemble of independent systems (i.e., not correlated with any other system in the Universe) which nevertheless requires a mixed density operator $\rho_A$; this is one of the non-mechanical states the existence of which was postulated to exist in Ref. [16] and is still controversial;

5. system A is a member of a heterogeneous ensemble $\{\{w, 1 - w\}, \{\rho_1, \rho_2\}\}$ where $0 < w < 1$ with $w\rho_1 + (1 - w)\rho_2 = \rho_A$, where $\rho_1$ and $\rho_2$ are any pair or combination of the above cases.

Are all these cases equivalent, in the sense that no (local) measurement on system A can possibly distinguish between them? If we exclude case 4 from the list, the current understanding is that the answer to this question is "yes".

Case 4 appears conceptually very different from Case 1, so much that one is tempted to reject it as unphysical. However, if there are no experimental ways to distinguish them, then we must consider them physically equivalent. The difference is only philosophical. If this is true, then saying that a system A is independent of any other system in the Universe is tantamount to saying that somewhere in the Universe there is a system B, entangled with A, so that AB is in a pure state. The mathematics say so, therefore it is conceivable that such a system B is there, somewhere in the Universe; but we have a simpler model of physical reality if we just assume (equivalently) the independence of system A, as done in Ref. [16]. As a matter of fact, there and in Ref. [15], it is assumed at the outset and explicitly that to be well-defined a state must be independent, much in the same way as, to be well defined, a system must be separable. Obviously, in cases when the system B, entangled with A, is not just hypothetical, but is part of the physical reality our model sets out to represent; then, we can show that, in the environment of A, there is an entangled B and so a better (or 'correct') model would include it, so as to consider system AB as independent, as so on.

But have experiments been performed to prove that the measurement statistics we produce by entanglement or by statistical mixing are indeed so entirely indistinguishable to be considered equivalent?

When it comes to entropy and its physical role in determining the adiabatic availability of a system, another challenging question arises. If the density operators represent homogeneous ensembles, what should be used to represent the heterogeneous ones [31]?

As shown in Refs. [16, 35], the expression for the thermodynamic entropy of a system in state $\rho$ is $S = -k_B \text{Tr}(\rho \ln \rho) = -k_B \sum_i p_i \ln(p_i/d_i)$ where $k_B$ is Boltzmann's constant, the $p_i$'s are the eigenvalues of $\rho$ and the $d_i$'s their respective degeneracies. Of course, $S = 0$ if the density operator $\rho$ is idempotent.

Consider the random statistical composition of two *different* preparations $\Pi_1$ and $\Pi_2$ each of which is homogeneous, i.e., it prepares every individual system in the ensemble in the individual state $\rho_1$ or $\rho_2$, respectively, each obtained, say, to fix ideas,

as in Case 1 above. The current notation[27] for this statistically mixed ensemble is $\{\{w, 1-w\}, \{\rho_1, \rho_2\}\}$ where $0 < w < 1$. Denote by $S_1$ and $S_2$ the entropies of the component states, and by $\Psi_1$ and $\Psi_2$ their adiabatic availabilities. It would seem logical that, since we are getting state $\rho_1$ a fraction $w$ of the times and $\rho_2$ the remaining ones, all properties should be weighted averages of the properties of the component states. So, for the ensemble we would expect quite naturally that $\langle S \rangle = w S_1 + (1-w) S_2$ and $\langle \Psi \rangle = w \Psi_1 + (1-w) \Psi_2$. Following von Neumann's quantum statistical recipe, instead, we should assign to this mixed ensemble the density operator $\rho_w = w \rho_1 + (1-w) \rho_2$, with the consequence that the corresponding values of $S$ and $\Psi$ are not what we expect. In fact, $S(\rho_w) > \langle S \rangle$ and $\Psi(\rho_w) < \langle \Psi \rangle$. It is as if the probability that the individual system belongs to one or the other component ensemble, mixes with the intrinsic probabilities of the individual systems, and does it so intimately and irrecoverably that the actual adiabatic availability is less than the average availability of the individual systems.

In terms of the battery example cited above, the paradox becomes more obvious. Say, $\rho_1$ is the state of the charged battery and $\rho_2$ that of the discharged, and suppose we have a reliable procedure $\lambda_1$ to extract the adiabatic availability $\Psi_1$ from the charged ones.[28] For example, we could connect the battery to a resistor and measure the electric current until it discharges. We do not know whether the next battery is charged or not, but by connecting a resistor to a discharged one we get no work out, so $\Psi_2 = 0$. Since the fraction of charged batteries is $w$, the overall net work output per battery will be $w \Psi_1$. Thus, we run into a ***paradox***, because we have just seen that according to von Neumann's recipe, we should get less than $w \Psi_1$.

# ENTROPY CHALLENGES FOR THE 21st CENTURY: IRREVERSIBILITY AND FUNDAMENTAL DYNAMICS

Theories that attempt to develop a fundamental first principle understanding of the origins of irreversibility face the requirement of compatibility with a wealth of objective experimental evidence. Fick's law of mass diffusion, Fourier's law heat conduction, Navier's law of momentum diffusion, Ohm's law of charge diffusion describe evidence that deeply entangles physics and engineering. In our view they describe bulk irreversible phenomena not surface phenomena, whereby the entropy generation occurs within each fluid or solid element by local spontaneous redistribution of mass, energy, momentum and charge. At steady state, the interactions with its neighbors maintain each fluid element in a state that otherwise would be nonequilibrium, by transferring out of

---

[27] The recent introduction of this notation is an implicit acknowledgment that the von Neumann recipe that would instead represent the heterogeneous ensemble by means of the mixed density operator $\rho = w \rho_1 + (1-w) \rho_2$ is misleading and incompatible, as we are about to see, with the measurement of observables that are nonlinear functionals of $\rho$, such as entropy and adiabatic availability.

[28] In quantum language, the procedure consists of devising a temporary change of the (parameters of the) Hamiltonian of the system for a precise time interval, and assume a unitary evolution generated by such Hamiltonian. When such change and time interval are tailor designed on a particular state $\rho$ of interest (such as the charged battery, or some given non-equilibrium state) we can make the system end in any other state with the same eigenvalues and degeneracies as the initial state $\rho$.

it exactly the entropy that it continuously generates by its bulk spontaneous tendency (attraction) toward a stable equilibrium internal distribution of mass, energy, momentum and charge. Such kind of irreversible phenomena are widely verified in all fields of science and engineering, where they are definitely perceived as real and objective, and occurring regardless of the state of knowledge of the observer or whether the system is being observed or not.

Yet, the most widely accepted physical explanations ultimately hinge on some "time averaging" or "coarse graining" argument justified by the "time scale of *observation*" being much longer than the correlation time between each particle and its surroundings. For example, the Boltzmann equation yields good predictions for not-too-dense gases, but all attempts to derive it from the reversible Hamilton-Liouville dynamics introduce at some point a key assumption (or *ad hoc* approximation) which basically "builds in" irreversibility. The assumption is that in the short time interval between one collision and the next, each particle "forgets" the correlations produced by the previous collision. Yet Hamilton-Liouville classical dynamics literally forbids such decoherence and spontaneous erasure of correlation (and so does the equivalent Schrödinger-Liouville-von Neumann unitary quantum dynamics). Thus, we are left with the following variant of the Loschmidt paradox: how come physical observations seem to agree with a model, like the Boltzmann classical equation (or the quantum Markovian master equation we discuss below), which "forces in" irreversibility by assuming a decoherence and erasure of correlation mechanism which clearly violates the fundamental dynamical law unconditionally assumed at the outset of the derivation to explain all phenomena?

Before discussing the prevailing model of irreversibility, let us make a few preliminary remarks about some little discussed important points and questions.

***The Second Law does not demand irreversibility. It is compatible with reversible dynamics, but not with unitary dynamics.***

The Second Law is often associated with irreversibility and "the arrow of time". However, in Ref. [16, p.450] it has been noted that the statement and the validity of the Second Law are independent of the existence of irreversibility, because the statement of existence and uniqueness of stable equilibrium states does not demand the existence of irreversible processes. In fact, all the results of Thermodynamics, including the fact that in most states not all the energy of the system can be extracted in the form of work, would maintain their importance and validity even in a scenario in which the Universe evolves reversibly remaining either at zero or at constant entropy. In other words, the Second Law holds valid and nontrivial irrespective of the Loschmidt paradox (1876).[29]

However, a little known entropy challenge emerges from the observation that if we assume that non-idempotent density operators of an isolated system evolve according to unitary Liouville-von Neumann quantum dynamics, $d\rho/dt = -i[H, \rho]/\hbar$, then the stable equilibrium states are many more than the (Hatsopoulos-Keenan statement of the) Second Law asserts. In fact, not only the maximal-entropy canonical density operators, but also all the other stationary density operators (such that $\rho H = H\rho$) are stable equilibrium [42]. This means that, since it contradicts it, unitary dynamics cannot possibly

---

[29] The Loschmidt paradox is [36]: how come we observe irreversible phenomena everyday and yet the fundamental laws of dynamics, that should explain all phenomena, are strictly reversible?

coexist with the Second Law (it violates the uniqueness of stable equilibrium for each given value of $\mathrm{Tr}(\rho H)$). Thus, considerations and derivations that assume the validity at the fundamental level of both unitary dynamics and the Second law, rest on faulty logic (contradictory premises). Adopting the assumption of unitary dynamics leaves for the Second Law only the non-fundamental (phenomenological) role (approximation) which it has according to the prevailing view. However, an alternative is possible, namely that the Second Law is fundamental and the dynamics is not unitary in general (we return on this in the last paragraph).

In a reversible Universe obeying unitary (Hamiltonian) dynamics, if a system is isolated, or adiabatic,[30] the (Liouville-von Neumann) equation of motion is $d\rho/dt = -i[H,\rho]/\hbar$, which of course reduces to the Schrödinger equation $d|\psi\rangle/dt = -iH|\psi\rangle/\hbar$ if $\rho$ is idempotent, $\rho = |\psi\rangle\langle\psi|$. As is well known, all the eigenvalues of $\rho$, their degeneracies and, therefore, the entropy are time invariant.

If by means of a (local) unitary process we try to extract the adiabatic availability of a system in a state $\rho_1$, the best we can do is to devise a temporary time-dependence of the Hamiltonian $H$ to be turned on only for a given time interval so that when we turn it off the system is in the state $\rho_2$ which commutes with $H$ and has eigenvalues arranged in decreasing order when represented with respect to the $H$-eigenvectors ordered by increasing energy eigenvalues.[31] In general the energy extracted in this way is less than or equal to the adiabatic availability. Except for very special $\rho_1$'s, to get out the full adiabatic availability requires a (locally) nonunitary process that changes the eigenvalues of $\rho$ at constant entropy, a task which according to standard Quantum Mechanics is possible only in principle because we need to have control not only on our system but also on all the systems with which ours is entangled.[32]

---

[30] We say that the system is *adiabatic* when it is separable (in the sense previously defined[19]) and its Hamiltonian operator may depend on some parameters $\boldsymbol{\beta}$ which describe the effects of external forces. If the parameters are time-independent, then the system is *isolated*.

[31] This important theorem, stated and proved in Ref. [16, pp.136-138], has been rediscovered as the main result in a recent paper [37] where the name "*ergotropy*" is coined to describe the maximum amount of energy that can be extracted in this way.

[32] From the perspective discussed above by which "local entropy" is due to entanglement between the local system $A$ and some other system $B$, we may always—in principle—extract the full adiabatic availability as follows. Based on the Schmidt decomposition theorem, we consider the maximally entangled "purifications" of the given local state $\rho_1$ and of the unique canonical local (stable equilibrium) state $\rho_{S1}$ with the same entropy as $\rho_1$. This means that $\rho_1$ and $\rho_{S1}$ are the reduced density operators of the pure states $|\psi_1\rangle\langle\psi_1|$ and $|\psi_{S1}\rangle\langle\psi_{S1}|$ of the overall entangled system $AB$. And since any pair of pure states can be interconnected by a unitary process, a suitable time-dependent Hamiltonian can be found which, when turned on for a suitable time interval and then turned off, can take one pure state to the other and hence change the state of $A$ from $\rho_1$ to $\rho_{S1}$, thus reversibly extracting the adiabatic availability of system $A$. Indeed, since initial and final pure states are maximally entangled, the "local entropy" of $B$ remains equal to that of $A$ and hence constant. The local process for $A$ is nonunitary. but of course the entropy of $AB$ remains zero. The challenges with this view are: (1) that of course we seldom have access and control of the entangled system $B$ which might involve many degrees of freedom and radiation that has traveled far away; (2) that there are infinite combinations of systems $B$ and pure states of $AB$ that all reduce to the same local states $\rho_1$ and $\rho_{S1}$ of $A$; (3) that "local entropies" are not additive for correlated systems (in fact the difference between sum of local entropies and entropy of the overall system is a measure of the degree of correlation of the subsystems) and therefore making a "local entropy" balance becomes tricky, if not impossible.

***Irreversibility does not demand the phase-space volume to increase.***

This remark follows from a somewhat technical property of the quantum entropy functional, $s(\rho) = -k_B \text{Tr}(\rho \ln \rho)$, which in the so-called *classical limit* $\hbar \to 0$ is usually claimed to tend to the classical functional[33] $s^{cl}(w) = -k_B \int \frac{dq\,dp}{2\pi\hbar} w(q,p) \ln w(q,p)$ where $w$ is either the Blokhintzev, the Wigner or the Wehrl phase-space distribution that map density operators to points in classical phase space (see [39]). In the limit as $\hbar \to 0$ the spectral expansion of the density operator $\rho$ induces [39] a partition of the classical phase space $\Omega$ into disjoint cells $\Omega_j$ each belonging to a distinct eigenvalue $p_j$ of $\rho$, and having a phase-space volume equal to the degeneracy $d_j$ of the corresponding eigenvalue $p_j$. In particular, the sum of the degeneracies of the nonzero eigenvalues is the *overall phase-space volume* occupied by the system in the given state (the volume of the support of $w$, i.e. where $w(q,p) \neq 0$). Therefore a Liouville-von Neumann unitary evolution of $\rho$ induces as $\hbar \to 0$ a Liouville evolution of $w$ which preserves both the *overall phase-space volume* and the value of the entropy. When it comes to describing irreversibility, the foregoing observations prompt an important and seldom recognized clarification: *conservation of phase-space volume* is *not* tantamount to *thermodynamic reversibility*. The entropy in the quantum description depends on both the $p_j$'s and the $d_j$'s; in the classical description it depends on the shape of $w$, not just the volume of its support. Therefore, to describe entropy changes, an evolution equation should capitalize on either or both of two independent and rather different mechanisms: (1) phase-space-volume-changing expansion, contraction or diffusion, and (2) constant-phase-space-volume redistribution.[34]

***Irreversibility does not demand dynamical map to be non-invertible.***

This remark is about the common assertion that unitary evolution is "reversible" because it always "has an inverse".[35] The theory of dynamical semi-groups has been developed around the idea that non-invertibility—the feature that distinguishes a semi-group from a group of dynamical evolutions [42]—is an indispensable feature of irreversibility; the equation of motion can be solved only forward in time, not backwards; *causality* is thus retained only in "weak form": future states can be predicted from the present state, but the past cannot be reconstructed from the present. To challenge this idea, we note that the existence of thermodynamic irreversibility is not incompatible with causality in

---

[33] In 1979 Wehrl noted that [38] "a rigorous proof of this is nowhere found in the literature". The problem of giving restrictive conditions that define a complete classical phase-space representation of quantum kinematics for systems with both a classical and a quantum description appears to be still unresolved, although heuristic arguments [39] do support the usual claim that $s(\rho) \to s^{cl}(w)$ in the classical limit $\hbar \to 0$.

[34] For example, the nonunitary part of the evolution equation proposed in Refs. [31, 40, 41] continuously redistributes the eigenvalues of $\rho$ while preserving its null eigenspace, which in the classical limit $\hbar \to 0$ means that it redistributes the shape of $w(q,p)$ while preserving its support.

[35] If $\rho(0)$ and $\rho(t)$ are the states at times 0 and $t$, they are related by the dynamical map $\rho(t) = \Lambda_t(\rho(0))$, i.e., the solution of the equation of motion for the time interval from 0 to $t$ with initial condition $\rho(0)$. If the inverse map exists, it points from the final state back to the initial state, $\rho(0) = \Lambda_t^{-1}(\rho(t))$. In other words, there is a one-to-one correspondence between initial and final state. Lack of inverse (*non-invertibility*) means instead that for example many initial states lead to the same final state; therefore, given the final state it is impossible to reconstruct which of the compatible initial states it came from.

the strong sense: it does not necessarily forbid the possibility to reconstruct the past from the present. In fact, Refs. [43, 44] show that it is possible to conceive a dynamical law—fully compatible with all thermodynamics principles—that entails irreversibility and yet generates a "strongly causal" group of dynamical evolutions, with inverse defined everywhere, unique trajectories through every state, fully defined both forward and backward in time, thus allowing full reconstruction of the past from the present. It is an example of an *invertible* dynamics which nevertheless is largely irreversible in that in forward time and for an adiabatic system it entails and describes entropy generation along the direction of steepest entropy ascent.

### *The prevailing view. Irreversibility from the Markovian approximation.*

The prevailing model of irreversibility, starts from unitary dynamics but assumes that no system is truly isolated, so that even an initial pure state becomes mixed due to increasing system-environment entanglement.

The system-environment entanglement builds up due to interactions according to the standard Liouville-von Neumann unitary dynamics of the overall system-environment composite. By tracing out all enviromental degrees of freedom and making the so-called Markovian approximation,[36] the overall unitary dynamics gives rise to a system's reduced dynamics which is nonunitary, linear, completely positive and generated by the celebrated Kossakowski-Sudarshan-Gorini-Lindblad (KSGL) quantum master equation.[37] [38] In this widely accepted model, the assumption of *erasure of correlations* is the sole mechanism responsible for "entropy generation" [47], but it appears contradic-

---

[36] In this phenomenological model [45] a system $A$ is assumed to be weakly coupled with a reservoir $R$, so that they can exchange energy via unitary evolution of the overall state $\rho_{AR}$. The reservoir $R$ is modeled as a collection of a large number of quantum systems (many degrees of freedom, e.g., the modes of the electromagnetic field). Because of the weak coupling, the unitary dynamics of $\rho_{AR}$ produces both an energy exchange and a build up of correlations between the system and the reservoir. However, justified only by some heuristic reasoning, a crucial *additional assumption* is injected in the derivation (Markovian "approximation"): that correlations *smear out* rapidly enough so as to maintain $A$ and $R$ effectively decorrelated not only initially, but at all times. When the reduced density operator of the system is time averaged (coarse grained) over a sufficiently long time interval, which is nevertheless still much shorter that the system's time scale of interest, the average correlations becomes negligible, and the state is assumed to effectively factor at all times. The model is phenomenological and basically charges the reservoir's complexity for the system's (apparent) losing quickly its memory of past interactions.

[37] As is well known, the KSGL quantum Markovian master equation has the form [46]

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H,\rho] + \frac{1}{2}\sum_j \left(2V_j^\dagger \rho V_j - \{V_j^\dagger V_j, \rho\}\right) , \tag{1}$$

where the $V_j$'s are some operators on $\mathscr{H}$ (each term within the summation, often written in the alternative form $[V_j, \rho V_j^\dagger] + [V_j \rho, V_j^\dagger]$, is obviously traceless). It has been used for a number of successful models of dissipative quantum dynamics of open subsystems.

[38] Operators $V_j$ in Eq. (1) are in general interpreted as creation and annihilation, or transition operators. For example, by choosing [31], $V_j = c_{rs}|r\rangle\langle s|$, where $c_{rs}$ are complex scalars and $|s\rangle$ eigenvectors of the Hamiltonian operator $H$, and defining the transition probabilities $w_{rs} = c_{rs}c_{rs}^*$, equation (1) becomes

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] + \sum_{rs} w_{rs} \left(|s\rangle\langle\rho\rangle\langle s| - \frac{1}{2}\{|s\rangle\langle s|, \rho\}\right), \tag{2}$$

tory with the assumed underlying reversible unitary dynamics. The situation is parallel to what is needed to "derive" the classical Boltzmann equation from the underlying reversible Hamilton-Liouville dynamics.

A problem with the Markovian assumption and the theory of completely positive linear dynamical semigroups, is that the KSGL master equation preserves positivity of $\rho$ only in forward time, not backwards. We already noted that such "non-invertibility" is not a necessary feature of thermodynamic irreversibility but is instead related to the principle of causality. By seeking irreversibility through non-invertibility are we ready to give up the principle of causality, a keystone of scientific thinking and philosophical reasoning?

Moreover, it is (philosophically) hard to understand how diffusion of mass, momentum, energy and charge, could find their justification in a "loss of *information* on the time scale of the observer leading to rapid decoherence from the entanglement which continuously builds up by weak coupling with environmental degrees of freedom". Is this the real reason for the "universal tendency in nature to the dissipation of mechanical energy" already recognized by Kelvin in 1852 [48]?

### *An alternative view. Irreversibility built in the fundamental microscopic dynamical laws.*

A possible alternative is to assume a fundamental non-unitary extension of standard Schrödinger unitary dynamics not contradicting the Second Law nor any of the successful results of pure-state quantum mechanics, but entailing an objective entropy increase for mixed states. We have shown in Refs. [31, 40, 44] that such an approach is possible based on a steepest entropy ascent, i.e. maximal entropy generation, nonlinear and non-unitary equation of motion which reduces to the Schrödinger equation for pure states.[39] The challenge with this approach is to ascertain if the intrinsic irreversibility it implies at the single particle (local, microscopic) level is experimentally verifiable, or else its mathematics must only be considered yet another phenomenological tool, at the same level as the quantum Markovian master equations which, as we have seen, are not free of their own challenges.

---

or, equivalently, for the *nm*-th matrix element of $\rho$ in the $H$ representation,

$$\frac{\mathrm{d}\rho_{nm}}{\mathrm{d}t} = -\frac{i}{\hbar}\rho_{nm}(E_n - E_m) + \delta_{nm}\sum_r w_{nr}\rho_{rr} - \rho_{nm}\frac{1}{2}\sum_r (w_{rn} + w_{rm}) \,, \tag{3}$$

which, for the *n*-th energy level occupation probability $p_n = \rho_{nn}$, is the celebrated Pauli master equation

$$\frac{\mathrm{d}p_n}{\mathrm{d}t} = \sum_r w_{nr}p_r - p_n\sum_r w_{rn} \,. \tag{4}$$

[39] In a 1985 Nature editorial, John Maddox defined this approach "An adventurous scheme which may end arguments about the arrow of time". Here we do not discuss it further because it is detailed and discussed elsewhere in this volume. See the articles by Bedeaux, Gyftopoulos, Beretta, Gheorghiu-Svirschevski, and Von Spakovsky.

# CONCLUSION

Professor Keenan's method in teaching and mentoring was by "asking questions". In this introductory paper, we hope to have honored his memory by reviewing and formulating in our own language the many fundamental "questions" that arise when the various facets of the Second Law of Thermodynamics, regarding Entropy, Reversibility and Irreversibility, are confronted with Classical and Quantum Mechanics. Of these questions many are old and well known, but a few are less known if not new, and deserve more attention. Some researchers will remain convinced that all these questions are already well resolved by the currently prevailing theories and interpretations. We instead believe they still constitute a formidable "challenge", that present and future generations of scientists and engineers ought to "meet".

# ACKNOWLEDGMENTS

# REFERENCES

1.  R.J.E. Clausius, "On the Motive Power of Heat, and on the Laws which Can Be Deduced from It for the Theory of Heat", *Annalen der Physik und Chemie* (ed. J.C. Poggendorff) **79**, 368-397, 500-524 (1850).
2.  N.L.S. Carnot, *Reflections on the Motive Power of Fire: And other Papers on the Second Law of Thermodynamics*, Dover edition, 2005 (*Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*, A Mallet-Bachelier, Paris, 1824).
3.  J.P. Joule, "On the Mechanical Equivalent of Heat and on the Constitution of Elastic Fluids", *B.A.A.S. Report* **II**, 21 (1848).
4.  W. Thomson (Baron Kelvin of Largs, William Thomson), "An Account of Carnot's Theory of the Motive Power of Heat, with Numerical Results, Deduced from Regnault's Experiments on Steam", *Transactions of the Royal Society of Edinburgh* **16**, 541 (1849).
5.  R.J.E. Clausius, "Uber verschiedene fur die Anwendung bequeme Formen der Hauptgleichung der mechanischen", *Annalen der Physik und Chemie* (ed. J.C. Poggendorff) **125**, 325-400 (1865).
6.  J.C. Maxwell, *Theory of Heat*, Appleton, London, 1871.
7.  J.H. Jeans, *The Dynamical Theory of Gases*, Cambridge University Press, 1904.
8.  L. Szilard, "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings," *Zeitschrift fur Physik* **53**, 840 (1929): English translation in *Behavioral Science* **9**, 301 (1964).
9.  L. Brillouin, *Science and Information Theory*, Academic Press, 1956.
10. C.E. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal* **27**, 379, 623 (1948).
11. R. Landauer, "Irreversibility and Heat Generation in the Computing Process", *IBM Journal of Research and Development* **5**, 183 (1961).
12. C.H. Bennett, "The Thermodynamics of Computation - A Review", *Int. J. Theor. Phys.* **21**, 905 (1982).
13. M. Planck, *Treatise on Thermodynamics*, Longmans, Green and Co., 1927.
14. G.N. Hatsopoulos and J.H. Keenan, *Principles of General Thermodynamics*, Wiley, 1965.
15. E.P. Gyftopoulos and G.P. Beretta, *Thermodynamics. Foundations and Applications*, Dover, 2005 (first edition, Macmillan, 1991).

16. G.N. Hatsopoulos and E.P. Gyftopoulos, "A Unified Theory of Mechanics and Thermodynamics", *Found. Phys.* **6**, 15, 127, 439, 561 (1976).

17. G.P. Beretta, "Axiomatic Definition of Entropy for Nonequilibrium States", proceedings of the IX Joint European Thermodynamics Conference, Saint-Étienne, France, June 2007, *Int. J. Thermodynamics*, in press (2008).

18. H. Margenau, *The Nature of Physical Reality*, McGraw-Hill, 1950.

19. W. Heisenberg, "Ueber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", *Zeitschrift fur Physik* **18**, 172 (1927).

20. J.C. Slater, *Introduction to Chemical Physics*, McGraw-Hill, 1939.

21. S. Watanabe, *La Deuxième Théorème de la Thermodynamique et la Mécanique Ondulatoire*, Herman et Cie, Paris, 1935.

22. P. Demers, "Le Second Principe et la Théorie des Quanta", *Can. J. Res.* **22**, 27 (1944); **23**, 47 (1945).

23. M. Ozawa, "Uncertainty relations for noise and disturbancein generalized quantum measurements", *Annals of Physics* **311**, 350 (2004).

24. J.L. Park and H. Margenau, "Simultaneous measurement in quantum theory", *Int. J. Theor. Phys.* **1**, 211 (1968).

25. R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, "Thermodynamics of Quantum Information Systems - Hamiltonian Description", *Open Systems and Information Dynamics* **11**, 205 (2004).

26. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, 1932 (*Mathematical Foundations of Quantum Mechanics*, Engl. transl. of the 1932 German edition by R.T. Beyer, Princeton University Press, 1955, pp. 295-346).

27. W. Band and J.L. Park, "The Empirical Determination of Quantum States", *Found. Phys.* **1**, 133 (1970), 211, 339 (1971). See also U. Fano, "Decription of States in Quantum Mechanics by Density Matrix and Operator Techniques", *Rev. Mod. Phys.* **29**, 74 (1957).

28. P.A.M. Dirac, *Principles of Quantum Mechanics*, Oxford, 1930. The bra-ket notation was introduced in P.A.M. Dirac, "A New Notation for Quantum Mechanics", *Proceedings of the Cambridge Philosophical Society* **35**, p. 416 (1939)

29. For the original references and a reevaluation of the accepted scenario for the development of the new quantum mechanics in the 1920s see Mara Beller, "Matrix Theory before Schrödinger: Philosophy, Problems, Consequences", *Isis*, **74**, 469 (1983).

30. R. Horodecki, M. Horodecki, and P. Horodecki, "Balance of Information in Bipartite Quantum-Communication Systems: Entanglement-Energy Analogy", *Phys. Rev. A* **63**, 022310 (2001).

31. G.P. Beretta, "On the General Equation of Motion of Quantum Thermodynamics and the Distinction between Quantal and Nonquantal Uncertainties," Sc.D. thesis, MIT, Cambridge, MA, 1981; arXiv:quant-ph/0509116.

32. E. Schrödinger, "Probability relations between separated systems", *Proceedings of the Cambridge Philosophical Society* **32**, 446 (1936). For a recent account, see also Ref. [33].

33. G.P. Beretta, "The Hatsopoulos-Gyftopoulos Resolution of the Schrödinger-Park Paradox about the Concept Of State in Quantum Statistical Mechanics", *Mod. Phys. Lett. A* **21**, 2799 (2006).

34. J.L. Park," The Nature of Quantum States", *American Journal of Physics* **36**, 211 (1968).

35. E.P. Gyftopoulos and E. Cubuku, "Entropy: Thermodynamic Definition and Quantum Expression", *Phys. Rev. E* **55**, 3851 (1997).

36. J. Loschmidt, "Uber den Zustand des Warmegleichgewichtes eines Systemes von Kirpern mit Rucksicht auf die Schwerkraft", *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien, Mathematisch-Naturwissenschaftlichen Classe* **73**, 128 (1876).

37. A.E. Allahverdyan, R. Balian, and Th.M. Nieuwenhuizen, "Maximal Work Extraction from Finite Quantum Systems", *Europhysics Letters* **67**, 565 (2004).

38. A. Wehrl, "On the Relation between Classical and Quantum-Mechanical Entropy", *Rep. Math. Phys.* **16**, 353 (1979). See also A. Wehrl, "General properties of entropy", *Rev. Mod. Phys.* **50**, 221 (1978); and "A Remark on the Concavity of Entropy", *Found. Phys.* **9**, 939 (1979).

39. G.P. Beretta, "On the Relation between Classical and Quantum-Thermodynamic Entropy", *J. Math. Phys.* **25**, 1507 (1984).

40. G.P. Beretta, E.P. Gyftopoulos, J.L. Park, and G.N. Hatsopoulos, "Quantum Thermodynamics. A New Equation of Motion for a Single Constituent of Matter", *Nuovo Cimento B* **82**, 169 (1984); G.P. Beretta, "Quantum Thermodynamics of Nonequilibrium. Onsager Reciprocity and Dispercsio-Dissipation Relations", *Found. Phys.*, **17**, 365 (1987) and references therein; see also G.P. Beretta,

"The Second Law of Thermodynamics from Locally Maximal Entropy Generation Dynamics", in the present volume.

41. S. Gheorghiu-Svirschevski, "Nonlinear Quantum Evolution with Maximal Entropy Production", *Phys. Rev. A* **63**, 022105; and "Addendum", *ibid*, 054102 (2001). See also G.P. Beretta, "Maximal-Entropy-Production-Rate Nonlinear Quantum Dynamics Compatible with Second Law, Reciprocity, Fluctuation-Dissipation, and Time-Energy Uncertainty Relations", arXiv:quant-ph/0112046 (2001).

42. G.P. Beretta, "A Theorem on Lyapunov Stability and a Conjecture on a Property of Entropy", *J. Math. Phys.* **27**, 305 (1986).

43. G.P. Beretta, "Nonlinear model dynamics for closed-system, constrained, maximal-entropy-generation relaxation by energy redistribution", *Phys. Rev. E* **73**, 026113 (2006).

44. G.P. Beretta, "Positive Nonlinear Dynamical Group Uniting Quantum Mechanics and Thermodynamics", arXiv:quant-ph/0612215 (2006). G.P. Beretta, "Well-behaved nonlinear evolution equation for steepest-entropy-ascent dissipative quantum dynamics", *Int. J. Quantum Information* **5**, 249 (2007).

45. W.H. Louisell, "Quantum Statistical Properties of Radiation", Wiley, 1973, Chapter 6.

46. A. Kossakowski, *Bull. Acad. Sci. Math.* **20**, 1021 (1972); A. Kossakowski, *Rep. Math. Phys.* **3**, 247 (1972); R.S. Ingarden and A. Kossakowski, *Ann. Phys. (N.Y.)* **89**, 451 (1975); G. Lindblad, *Comm. Math. Phys.* **48**, 119 (1976).

47. V. Gorini, A. Kossakowski, and E.C.G. Sudarshan, *J. Math. Phys.* **17**, 821 (1976).

48. W. Thomson, "On a Universal Tendency in Nature to the Dissipation of Mechanical Energy", *Proceedings of the Royal Society of Edinburgh* **3**, 139 (1852).

49. J. Maddox, "Uniting Mechanics and Statistics", *Nature* **316**, 11 (1985).