

Quantum Solution to the Arrow-of-Time Dilemma

Lorenzo Maccone*

QUIT, Dip. A. Volta, 27100 Pavia, and Institute for Scientific Interchange, 10133 Torino, Italy
(Received 8 February 2008; revised manuscript received 3 September 2008; published 17 August 2009)

The arrow-of-time dilemma states that the laws of physics are invariant for time inversion, whereas the familiar phenomena we see everyday are not (i.e., entropy increases). I show that, within a quantum mechanical framework, all phenomena which leave a trail of information behind (and hence can be studied by physics) are those where entropy necessarily increases or remains constant. All phenomena where the entropy decreases must not leave any information of their having happened. This situation is completely indistinguishable from their not having happened at all. In the light of this observation, the second law of thermodynamics is reduced to a mere tautology: physics cannot study those processes where entropy has decreased, even if they were commonplace.

DOI: 10.1103/PhysRevLett.103.080401

PACS numbers: 03.65.Ta, 03.65.Ud, 03.67.-a, 05.70.Ce

Paradoxes have always been very fruitful in stimulating advances in physics. One which still lacks a satisfactory explanation is the Loschmidt paradox [1]. Namely, how can we obtain irreversible phenomena from reversible time-symmetric physical laws [2]? The irreversibility in physics is summarized by the second law of thermodynamics: entropy, which measures the degradation of the usable energy in a system, never decreases in isolated systems. Many approaches have been proposed to solve this conundrum, but most ultimately resort to postulating low entropy initial states (see, e.g., [3,4]), which is clearly an *ad hoc* assumption [5]. Others suggest that the thermodynamic arrow of time is in some way connected to the cosmological one [6], that physical laws must be modified to embed irreversibility [7], that irreversibility arises from decoherence [8], or from some time-symmetric mechanism embedded in quantum mechanics [9], etc. Recent reviews on this problem are given in Ref. [10].

Here I propose a different approach, based on existing laws of physics (quantum mechanics). I show that entropy in a system can both increase and decrease (as time reversal dictates), but that all entropy-decreasing transformations cannot leave *any* trace of their having happened. Since no information on them exists, this is indistinguishable from the situation in which such transformations do not happen at all: “The past exists only insofar as it is recorded in the present” [11]. Then the second law is forcefully valid: the only physical evolutions we see in our past, and which can then be studied, are those where entropy has not decreased.

I start by briefly relating the thermodynamic entropy with the von Neumann entropy, and introducing the second law. I then present two thought experiments, where entropy is deleted together with all records of the entropy-increasing processes: even though at some time the entropy of the system had definitely increased, afterward it is decreased again, but none of the observers can be aware of it. I conclude with a general derivation through the analysis of the entropy transfers that take place in physical transformations.

Entropy and the second law.—Thermodynamic entropy is a quantity that measures how the usable energy in a physical process is degraded into heat. It can be introduced in many ways from different axiomatizations of thermodynamics. The von Neumann entropy of a quantum system in the state ρ is defined as $S(\rho) \equiv -\text{Tr}[\rho \log_2 \rho]$. When applicable, these two entropies coincide (except for an inconsequential multiplicative factor). This derives from an argument introduced by Einstein and extended by Peres [12] (e.g., both the canonical and the microcanonical ensemble can be derived from quantum mechanical considerations [13,14]). For our purposes, however, it is sufficient to observe that thermodynamic and von Neumann entropies can be interconverted, employing Maxwell-demons [15,16] or Szilard-engines [17,18]: useful work can be extracted from a single thermal reservoir by increasing the von Neumann entropy of a memory space.

There are many different formulations of the second law, but we can summarize them by stating that, in any process in which an isolated system goes from one state to another, its thermodynamic entropy does not decrease [19]. There is a hidden assumption in this statement. Whenever an isolated system is obtained by joining two previously isolated systems, then the second law is valid only if the two systems are initially uncorrelated, i.e., if their initial joint entropy is the sum of their individual entropies. It is generally impossible to exclude that two systems might be correlated in some unknown way and there is no operative method to determine whether a system is uncorrelated from all others. Thus, in thermodynamics all systems are considered uncorrelated, unless it is known otherwise. Without this assumption, it would be impossible to assign an entropy to any system unless the state of the whole universe is known: a normal observer is limited in the information she can acquire and on the control she can apply. This implies that thermodynamic entropy is a subjective quantity, even though *for all practical situations* this is completely irrelevant: the eventual correlations in all macroscopic systems are practically impossible to control and exploit. Even

though they are ignored by the normal observer, correlations between herself and other systems do exist. Until they are eliminated, the other systems cannot decrease their entropy. A physical process may either reduce or increase these correlations. When they are reduced, this may seem to entail a diminishing of the entropy, but the observer will not be aware of it as her memories are correlations and will have been erased by necessity (each bit of memory is 1 bit of correlation and, until her memory has been erased, the correlations are not eliminated). Instead, when the physical process increases these correlations, she will see it as an increase in entropy. The observer will then only be aware of entropy nondecreasing processes.

The above analysis is limited to systems that are somehow correlated with the observer. One might then expect that she could witness entropy-decreasing processes in systems that are completely factorized from her. That is indeed the case: statistical microscopic fluctuations can occasionally decrease the entropy of a system (the second law has only a statistical valence). However, an observer is macroscopic by definition, and all remotely interacting macroscopic systems become correlated very rapidly (e.g. Borel famously calculated that moving a gram of material on the star Sirius by 1 m can influence the trajectories of the particles in a gas on earth on a time scale of μs [20]). This is the same mechanism at the basis of quantum decoherence [8], and it entails that in practice the above analysis applies to all situations: no entropy decrease in macroscopic systems is ever observed.

In what follows I will make these ideas rigorous.

Thought experiments.—The quantum information theory mantra “information is physical” [21] implies that any record [22] of an occurred event can be decorrelated from such event by an appropriate physical interaction. If *all* the records of an event are decorrelated from it, then by definition there is no way to know whether this event has ever happened. This situation is indistinguishable from its not having happened. If this event has increased the entropy, the subsequent erasing of all records can (*will*) produce an entropy decrease without violation of any physical law. We now analyze two such situations, an imperfect transmission of energy and a quantum measurement.

Alice’s lab is perfectly isolated, so that to an outside observer (Bob), its quantum evolution is unitary. Analyze the situation in which Bob sends Alice some energy in the form of light, a multimode electromagnetic field in a zero-entropy pure state. We suppose that, to secure the energy Bob is sending her, she uses many detectors which are not matched to his modes. Given a system in almost any possible pure state, all its subsystems which are small enough are approximately in the canonical state [13]. This implies that, if each of Alice’s detectors is sensitive to only a small part of Bob’s modes, the detectors mostly see thermal radiation, and she feels them warming up. She will then be justified in assigning a nonzero thermodynamic entropy to her detectors, as she sees them basically

as thermal-equilibrium systems. One might object that she is mistaken, since the states of the detectors are not uncorrelated. However, since she ignores the correlations, she cannot use such correlations to extract energy from the detectors. Alice concludes that most of the energy Bob sent her has been wasted as heat, raising the thermodynamic entropy of her lab. Suppose now that Bob has complete control of all the degrees of freedom in her lab. He knows and can exploit the correlations to recover all the energy he had initially given Alice. Of course, although possible in principle, he needs a dauntingly complex transformation, which requires him to be able to control a huge number of her lab’s degrees of freedom (including the brain cells where her memories are, and the notepads where she wrote the temperatures). To extract the energy, since it was initially locked in a pure state of the field, he must return it to a system in a zero-entropy pure state, i.e., factorized from all the other degrees of freedom of Alice’s lab. Then he must erase *all* the correlations between them: at the end of Bob’s recovery, Alice cannot remember feeling her detectors warm up, they are cool again, her notepads contain no information, and all the energy initially in the electromagnetic field is again available, even though (from Alice’s point of view) most of it was definitely locked into thermal energy at one time.

The second thought experiment [23] is a prototypical quantum measurement. Bob prepares a spin-1/2 particle oriented along the x axis, e.g., in a spin $|\rightarrow\rangle$ state and hands it to Alice. She sends it through a Stern-Gerlach apparatus oriented along the z axis [12]. The measurement consists in coupling the quantum system with some macroscopic degrees of freedom (a reservoir), not all of which are under the control of the experimenter [24], whence the irreversibility. Notice that, $|\rightarrow\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$, where $|\uparrow\rangle$ and $|\downarrow\rangle$ are the eigenstates of a z measurement operator. Hence, this apparatus will increase the entropy of the spin system by 1 bit [15]: Before the readout, the spin state will be in the maximally mixed state $(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)/2$. After Alice has looked at the result, she has transferred this 1 bit of entropy to her memory. From the point of view of Bob, outside her isolated lab, Alice’s measurement is simply a (quantum) correlation of her measurement apparatus to the spin. The initial state of the spin $|\rightarrow\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ evolves into the correlated (entangled) state

$$(|\uparrow\rangle|_{\text{Alice sees “up”}} + |\downarrow\rangle|_{\text{Alice sees “down”}})/\sqrt{2}, \quad (1)$$

where the first ket in the two terms refers to the spin state, whereas the second ket refers to the rest of Alice’s lab. Thus, from the point of view of Bob, Alice’s measurement is an evolution similar to a controlled-NOT unitary transformation of the type $U_{\text{cnot}}(|0\rangle + |1\rangle)|0\rangle = |0\rangle|0\rangle + |1\rangle|1\rangle$. Such a transformation can be easily inverted, as it is its own inverse. Analogously, Bob can flip a switch and invert Alice’s measurement. At the end of his operation, all records of the measurement result (Alice’s brain cells, the apparatus gauges) will have been decorrelated from

the spin state. She will remember having performed the measurement, but she will be (*must* be) unable to recall what the measurement result was and the spin is returned to a pure state. Bob's transformation is not necessarily a reversion of the dynamics of Alice's lab.

In both the above experiments, from Alice's point of view, entropy definitely has been created after she has interacted with Bob's light or his spin. However, this entropy is, subsequently, coherently erased by Bob. At the end of the process, looking back at the evolution in her lab, she cannot see any violation of the second law: she has no (cannot have any) record of the fact that entropy at one point had increased.

Entropic considerations.—The above thought experiments exemplify a general situation: entropy can decrease, but its decrease is accompanied by an erasure of any memory that the entropy-decreasing transformation has occurred. In fact, any interaction between an observer A and a system C which decreases their entropy by a certain quantity, must also reduce their quantum mutual information by the same amount (unless, of course, the entropy is dumped into a reservoir R). The quantum mutual information $S(A:C) \equiv S(\rho_A) + S(\rho_C) - S(\rho_{AC})$ measures the amount of shared quantum correlations between the two systems A and C (ρ_{AC} being the state of the system AC , and ρ_A and ρ_C its partial traces, i.e., the states of A and C).

Taking the cue from [25], I now prove the above assertion, namely, I show that

$$\Delta S(A) + \Delta S(C) - \Delta S(R) - \Delta S(A:C) = 0, \quad (2)$$

where $\Delta S(X) \equiv S_t(\rho_X) - S_0(\rho_X)$ is the entropy difference between the final state at time t and the initial state of the system X , and where $\Delta S(A:C) = S_t(A:C) - S_0(A:C)$ is the quantum mutual information difference. Choose the reservoir R so that the joint state of the systems ACR is pure and so that the evolution maintains the purity (R is a purification space). Then the initial and final entropies are $S_0(AC) = S_0(R)$ and $S_t(AC) = S_t(R)$, respectively. Thus we find $S_0(AB) = S_t(AB) - \Delta S(R)$ which, when substituted into the left-hand-side term of (2), shows that this term is null. [This proof is valid also if the evolution is not perfectly known, i.e., if it is given by a random unitary map.]

A memory of an event is a physical system A which has nonzero classical mutual information on a system C that bears the consequences of that event. Then, the erasure of the memory follows from an elimination of the quantum mutual information $S(A:C)$ if this last quantity is an upper bound to the classical mutual information $I(A:C)$. Thus, we must show that for any POVM measurement $\{\Pi_i^{(a)} \otimes \Pi_j^{(c)}\}$ extracting information separately from the two systems ($\Pi_i^{(a)}$ acting on A and $\Pi_j^{(c)}$ on C),

$$S(A:C) \geq I(A:C), \quad (3)$$

where $I(A:C)$ is the mutual information of the POVM's measurement results. A simple proof of this statement

exists (e.g. see [26,27]): use the equality $S(A:C) = S(\rho_{AC} || \rho_A \otimes \rho_C)$, where $S(\rho || \sigma) \equiv \text{Tr}[\rho \log_2 \rho - \rho \log_2 \sigma]$ is the quantum relative entropy. This quantity is monotonic for application of CP maps [26], i.e., $S(\rho || \sigma) \geq S(\mathcal{N}[\rho] || \mathcal{N}[\sigma])$ for any transformation \mathcal{N} that can be written as $\mathcal{N}[\rho] = \sum_k A_k \rho A_k^\dagger$, with the Kraus operators A_k satisfying $\sum_k A_k^\dagger A_k = 1$. Consider the ‘‘measure and re-prepare’’ channel, i.e., the transformation $\mathcal{N}[\rho] = \sum_n \text{Tr}[\Pi_n \rho] |n\rangle\langle n|$ where $\{|n\rangle\}$ is a basis, and Π_n is a POVM element (i.e. a positive operator such that $\sum_n \Pi_n = 1$). It is a CP map, since it has a Kraus form

$$A_{nm} = |n\rangle\langle v_m^{(n)}| \sqrt{p_m^{(n)}}, \quad \text{with} \quad \Pi_n = \sum_m p_m^{(n)} |v_m^{(n)}\rangle\langle v_m^{(n)}|.$$

Using the monotonicity of the relative entropy under the action of the map \mathcal{N} , we find

$$\begin{aligned} S(A:C) &= S(\rho_{AC} || \rho_A \otimes \rho_C) \geq S(\mathcal{N}[\rho_{AC}] || \mathcal{N}[\rho_A \otimes \rho_C]) \\ &= \sum_{ij} p_{ij} \log_2 p_{ij} - \sum_{ij} p_{ij} \log_2 (q_i r_j) = I(A:C), \end{aligned}$$

where $p_{ij} \equiv \text{Tr}[\Pi_i^{(a)} \otimes \Pi_j^{(c)} \rho_{AC}]$, $q_i \equiv \text{Tr}[\Pi_i^{(a)} \rho_A]$, and $r_j \equiv \text{Tr}[\Pi_j^{(c)} \rho_C]$.

The interpretation of Eq. (2) is that, if we want to decrease the entropy of the system C (somehow correlated with the observer A) without increasing the entropy of a reservoir R , we need to reduce the quantum mutual information between C and the observer A . The fact that mutual information can be used to decrease entropy was already pointed out by Lloyd [25] and Zurek [28].

The implications of the above analysis can be seen explicitly by employing Eq. (2) twice, by considering an intermediate time when $S(C)$ is higher than at the initial and final times. The entropy $S(C)$ of the system is high at the intermediate time after an entropy-increasing transformation, and then (if no entropy-absorbing reservoir R is used) it can be reduced by a successive entropy-decreasing transformation at the cost of reducing the mutual information between the observer and C . Even though the entropy $S(C)$ (as measured from the point of view of the observer A) does decrease, the observer is not aware of it, as the entropy-decreasing transformation must factorize her from the system C containing information on the prior entropy-increasing event: her memories of such event must be part of the destroyed correlations.

What we have seen up to now is that any decrease in entropy of a system that is correlated with an observer entails a memory erasure of said observer. That might seem to imply that an observer should be able to see entropy-decreasing processes when considering systems that are uncorrelated from her. In fact, at microscopic level, statistical fluctuations do decrease occasionally the entropy. However, the correlations between any two macroscopic systems build up continuously, and at amazing rates [20]: this is how decoherence arises [8]. Then no observer is really factorized with respect to any macroscopic system

she observes and entropy decreases of a macroscopic system becomes unobservable (unless extreme care is taken to shield the system under analysis). Only microscopic systems can be considered factorized from an observer for a period of time long enough to see entropy decrease from fluctuations.

Conclusions.—In this Letter I gave a quantum solution to the Loschmidt paradox, showing that all physical transformations where entropy is decreased cannot relinquish any memory of their having happened from the point of view of any observer: both normal observers that interact with the studied systems and external superobservers that keep track of all the correlations. Thus they are irrelevant to physics. Quantum mechanics is necessary to this argument. In the above derivation, we have used the property that the entropy of a joint system can be smaller than that of each of its subsystems. This is true of von Neumann entropy, but not true if entropy is calculated using classical probability theory: then the entropy of a joint system is always larger than that of its subsystem with largest entropy.

In a quantum cosmological setting, the above approach easily fits in the hypothesis that the quantum state of the whole Universe is a pure (i.e., zero entropy) state evolving unitarily (e.g., see [29,30]). One of the most puzzling aspects of our Universe is the fact that its initial state had entropy so much lower than we see today, making the initial state highly unlikely [4]. Joining the above hypothesis of a zero-entropy pure state of the universe with the second law considerations analyzed in this Letter, it is clear that such puzzle can be resolved. The universe may be in a zero-entropy state, even though it appears (to us, internal observers) to possess a higher entropy. However, it is clear that this approach does not require dealing with the quantum state of the whole Universe, but it applies also to arbitrary physical systems.

I thank S. Lloyd, G. M. D'Ariano, V. Giovannetti, G. P. Beretta, and A. Winter for discussions and feedback.

*Present address: RLE, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

- [1] J. Loschmidt, Sitzungsber. Kais. Akad. Wiss. Wien, Math. Naturwiss. Classe, II. Abteilung **73**, 128 (1876).
- [2] More precisely: from laws symmetric for CPT-inversion.
- [3] R. Penrose, in *General Relativity, an Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979).
- [4] R. M. Wald, arXiv:gr-qc/0507094v1.
- [5] E. G. D. Cohen and T. H. Berlin, *Physica* (Amsterdam) **26**, 717 (1960).
- [6] T. Gold, *The Nature of Time* (Cornell University Press, Ithaca, 1967); P.C.W. Davies, *The Physics of Time Asymmetry* (Surrey University Press, London, 1974); S. M. Carroll and J. Chen, *Gen. Relativ. Gravit.* **37**, 1671 (2005).
- [7] G. C. Ghirardi, A. Rimini, and T. Weber, *Phys. Rev. D* **34**, 470 (1986); D.Z. Albert, *Quantum Mechanics and Experience* (Harvard University Press, Cambridge, MA, 1994); E. P. Gyftopoulos and G. P. Beretta, *Thermodynamics: Foundations and Applications* (Dover, London, 2005).
- [8] H. D. Zeh, *The Physical Basis of the Direction of Time* (Springer-Verlag, Berlin, 1999).
- [9] L. S. Schulman, *Time's Arrow and Quantum Measurement* (Cambridge University Press, Cambridge, England, 1997).
- [10] *Physical Origins of Time Asymmetry*, edited by J. J. Halliwell, J. Pérez-Mercader, and W. H. Zurek (Cambridge University Press, Cambridge, England, 1996); H. Price, *Time's Arrow and Archimedes' Point* (Oxford University Press, Oxford, 1996); *Time's Arrows Today*, edited by S. F. Savitt (Cambridge University Press, Cambridge, England, 1997).
- [11] J. A. Wheeler, as quoted in W. H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003).
- [12] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Dordrecht, 1993).
- [13] S. Popescu, A. J. Short, and A. Winter, *Nature Phys.* **2**, 754 (2006).
- [14] J. Gemmer, A. Otte, and G. Mahler, *Phys. Rev. Lett.* **86**, 1927 (2001); H. Tasaki, *Phys. Rev. Lett.* **80**, 1373 (1998).
- [15] S. Lloyd, *Phys. Rev. A* **56**, 3374 (1997).
- [16] *Maxwell's Demon: Entropy, Information, Computing*, edited by H. S. Leff and A. F. Rex (Princeton University Press, Princeton, 1991); C. H. Bennett, *Int. J. Theor. Phys.* **21**, 905 (1982); V. Serreli, C.-F. Lee, E. R. Kay, and D. A. Leigh, *Nature* (London) **445**, 523 (2007).
- [17] L. Szilard, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, 1983), p. 539.
- [18] M. O. Scully, *Phys. Rev. Lett.* **87**, 220601 (2001).
- [19] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, Boston, 1965).
- [20] E. Borel, *Le Hasard* (Alcan, Paris, 1914).
- [21] R. Landauer, *Proceedings of the Workshop on Physics and Computation Phys Comp'92* (IEEE Press, Los Alamitos, 1992), p. 1.
- [22] By record, I consider any physical system which is correlated with the event, such as its description on the pages of a book, or the neural pattern in a witness's brain.
- [23] *The Many-Worlds Interpretation of Quantum Mechanics*, edited by B. S. DeWitt and R. N. Graham (Princeton University Press, Princeton, 1973).
- [24] A. Peres, *Phys. Rev. D* **22**, 879 (1980).
- [25] S. Lloyd, *Phys. Rev. A* **39**, 5378 (1989).
- [26] H. P. Yuen and M. Ozawa, *Phys. Rev. Lett.* **70**, 363 (1993).
- [27] D. DiVincenzo *et al.*, *Phys. Rev. Lett.* **92**, 067902 (2004).
- [28] W. H. Zurek, in *Frontiers of Nonequilibrium Statistical Physics*, edited by G. T. Moore and M. O. Scully (Plenum Press, New York, 1984), p. 151.
- [29] M. Tegmark, *Found. Phys. Lett.* **9**, 25 (1996).
- [30] D. Z. Albert, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 2, p. 127 (University of Chicago Press, Chicago, 1988).