

THE CRUCIAL ROLE OF THERMODYNAMIC GATES IN LIVING SYSTEMS

David W. Snoke*

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Abstract

Living systems can be characterized as seeming to violate the second law of thermodynamics in their operation, but not actually doing so, due to clever design. This type of operation raises questions of how these processes could be started. In this paper, I discuss how all forms of machines that have this type of operation rely crucially on the existence of “gates,” that is, interfaces between different populations that can either switch on or switch off entropic flow. These gates are directly comparable to transistors in modern electronic circuits. This allows analysis of biological systems to map directly to that of electronic and other engineered systems.

Keywords

second law of thermodynamics • nonequilibrium quantum mechanics • origin of life • definition of life • Maxwell's demon

1 Introduction

As discussed in a previous paper [1], a living system intrinsically behaves like a “Maxwell's demon” that sorts atoms in a gas. Several conclusions follow from this. First, if the demon is ignored, then the entropy of the gas appears to go down, in violation of the second law of thermodynamics. Second, if the Maxwell's demon is accounted for as part of the system along with the gas, but not the external environment to which the demon may vent heat, then the entropy in the gas+demon system will not violate the second law *only* if the demon is assumed to have low entropy at the start of the process. Third, if the entire system of gas+demon+environment is taken into account, then the total entropy of the whole system always goes up, because the demon must vent some heat to operate. This last statement is well known as a result of the Landauer argument [2].

While the first and third statements are well known, the second one is less well appreciated. It amounts to saying that not only must the total entropy in a whole system increase, but also the entropy in any sufficiently large subsystem must also increase. For example, in the case of a gas with two isotopes of the same element in a gas, the gas as a whole cannot move away from equilibrium by trading entropy between the two isotopes such that entropy of one isotope

goes down while the entropy in the other increases [3]. In the same way, the demon+gas subsystem cannot decrease its total entropy, no matter what the environment does. Therefore it must have started with low entropy. Since entropy is defined in terms of the total number of equivalent states (for a review of the definition of entropy, see, e.g., Ref. [1]), the only way the initial state could have lower entropy than the final state is if the microscopic states of the gas were not equivalent, as detected by the demon. To an outside observer, the entropy appears to decrease because many states are judged as equivalent, but to the demon, those states are distinguishable, and therefore the entropy is low even before the demon does any sorting. This analysis need not involve any philosophy of knowledge and consciousness, if we define inequivalent states as having different macroscopic results. Here, those are the actions of the demon to respond to information about the states of the gas. Thus the “demon” could be an inanimate machine, such as a robotic system.

Given this, we can ask what minimum characteristics are needed for a Maxwell's demon. It was argued in Ref. [1] that every “machine” is a type of Maxwell's demon, having the functions of detecting some information about the

* Corresponding author e-mail: snoke@pitt.edu

environment and producing some macroscopic action based on this information, which makes some states of the system nonequivalent. Simple machines process just a few bits of information, while complex machines and living systems process many more bits of information, reducing the number of equivalent states even further.

In this paper, I analyze the essential characteristics of “machines” from the standpoint of thermodynamics (in particular, using the results of quantum thermodynamics). This will lead to a unified description of spontaneous emergence, simple and complex machines, and life. We will see that switches or gates with transistor-like action are essential elements.

2 Review of why subsystems are subject to the second law of thermodynamics

2.1 Summary of quantum Boltzmann results

The discussion in Section 1 took it as a given that the entropy in any sufficiently large subsystem must always increase. This could be viewed as a statement of probability, which could be violated in unlikely cases. However, modern quantum mechanics makes a much stronger statement, that the second law is deterministically true in any system that does not have specially rigged (“fine-tuned”) initial conditions. Section 2.2 reviews the quantum mechanical calculation that leads to this conclusion. Since the proof involves a fair degree of mathematics, I summarize here the results, for those readers who wish to skip that section. The main results are

- The second law of thermodynamics is not a statistical law, that is, not merely a statement of high probability, but is a deterministic result of the time evolution of the quantum wave function in a system with many degrees of freedom.
- Irreversibility and increase of entropy always occur in an sufficiently large, ergodically connected, closed system unless the initial conditions of the system are fine tuned to a fantastic degree. “Sufficiently large” here means that the Poincaré recurrence time is much longer than any other timescale of the system; this is typically the case even for relatively small systems of 100 particles or so, unless they are highly constrained in their motion. (See Section 2.4 for a discussion of the Poincaré recurrence theorem.) “Ergodically connected” here means that any given state of the system is connected dynamically to all the other states of the system; no states are walled off. “Closed” does not mean that the system is finite in size; it means that all of the interactions of the system are accounted for by internal processes.

- The quantum Boltzmann equation gives a natural time scale for equilibration, which can be called the thermalization time, in any system. Transients and fluctuations might give a local decrease of entropy for short times, but these are damped out on time scales long compared to the intrinsic thermalization time.

In addition, in Section 2.3, I show the following:

- The entropy not only of a whole system, but of every sufficiently large, ergodically connected subset of a system always obeys the second law of thermodynamics.
- Finally, while it is not strictly proven, numerical solutions of the quantum Boltzmann equation indicate that even when a system does not satisfy the requirements above (large, ergodically connected, and closed), to the degree that the system approximates a large, ergodically connected, closed system, to that same degree the behavior of the system is well approximated by the second law of thermodynamics.

2.2 Quantum mechanical derivation of the second law

The proof summarized below was originally presented in Ref. [5]. To start, we need to introduce the notation for a complicated system with many degrees of freedom. In quantum mechanics, we can write the full wave function of a whole system in terms of “Fock states,” which give the amplitude of the wave function in each of the substates that define the system, which define the whole range of possible states of the system. (Mathematically, these are collectively called a “complete set of orthonormal states.”) These Fock states are written as

$$|\psi\rangle = |N_1, N_2, N_3, \dots\rangle, \quad (1)$$

where N_1, N_2 etc. are integers that represent the number of quanta in each substate. These quanta can be interpreted as “particles,” but need not be; instead they can be taken to represent simply the strength, or spectral weight, of the wave in each of its quantum resonances (see Ref. [4], Chapters 1 and 2, for more discussion of this). In a large system, there can be many, many possible Fock states; for example, in typical solids or gases there are around 10^{20} particles in a volume about a millimeter in size, and all possible permutations of all possible different numbers of these in different substates must be accounted for.

Because quantum mechanics allows superpositions of many wave states, the most general form of the full wave function of a system has the form

$$|\psi\rangle = \alpha |N_1, N_2, N_3, \dots\rangle + \beta |N'_1, N'_2, N'_3, \dots\rangle + \dots, \quad (2)$$

where the complex numbers $\alpha, \beta \dots$ give the amplitudes of each of the individual Fock states, and the sum includes all possible Fock states with all possible integer values of $N_1, N_2, \dots$, subject to the constraint $\sqrt{|\alpha|^2 + |\beta|^2 + \dots} = 1$. We can write this compactly as the sum

$$|\psi\rangle = \sum_i \alpha_i |N_1^{(i)}, N_2^{(i)}, N_3^{(i)}, \dots\rangle, \quad (3)$$

where i sums over all possible distinct Fock states. For a many-particle system, this wave function is quite complicated, with many values of α_i for all possible Fock states, which are formed for all possible values of N in each possible substate.

The information in this large and complicated wave state is typically analyzed via “correlation functions.” The simplest form of these correlation functions is the set of all possible correlations of two states. We write these in terms of the operator

$$\hat{\rho}_{nm} = a_n^\dagger a_m, \quad (4)$$

where a_m means “for each Fock state in the sum, multiply by $\sqrt{N_m}$ and reduce N_m by one,” and $a_n^\dagger$ means, “for each Fock state in the sum, multiply by $\sqrt{1 \pm N_n}$ and increase N_n by one.” The $+$ in this last formula is for bosons (integer-spin quanta) and the $-$ is for fermions (half-integer-spin quanta). The $a_n^\dagger$ and a_n operators are typically called “creation” and “destruction” operators, respectively, in the quantum algebra, since they change the number of quanta in each Fock state.

We can then talk of the “density matrix” as the set of all possible correlations. The square root rules given above ensure that the “diagonal” terms of the matrix are those that give the average number of quanta (“particles”) in each individual state:

$$\begin{aligned} \langle \hat{\rho}_{nn} \rangle &= \langle N_n \rangle = \langle \psi | a_n^\dagger a_n | \psi \rangle \\ &= \sum_i |\alpha_i|^2 N_n^{(i)}. \end{aligned} \quad (5)$$

This is always a real number, but it need not be an integer; fractional values are allowed because the system can be in a superposition of different numbers of particles in any given resonant state. The “off-diagonal” terms of the density matrix define the wave correlation between different states:

$$\begin{aligned} \langle \hat{\rho}_{nm} \rangle &= \langle \psi | a_n^\dagger a_m | \psi \rangle \\ &= \sum_{i,j} \alpha_i^* \alpha_j \sqrt{(N_n^{(i)} \pm 1) N_m^{(j)}}. \end{aligned} \quad (6)$$

where α_i^* is the complex conjugate of α_i . Since α_i^* and α_j are complex numbers, their product is in general also complex, which means that $\langle \hat{\rho}_{nm} \rangle$ can be written as $A e^{i\theta}$, where A is a real-valued amplitude, and θ is a “phase factor” that ranges

from 0 to 2π . The off-diagonal terms of the density matrix are therefore often called measures of the “phase coherence.”

The time evolution of the elements of the density matrix can be computed via deterministic evolution of the many-body wave function; the equation that gives the time evolution of the diagonal elements, in the limit of strong decoherence, is known as the *quantum Boltzmann equation*. The quantum Boltzmann equation has different forms for different types of interactions, but has the same overall irreversible behavior for all types of interactions. Ref. [5] showed that for collisional interactions, it has the form

$$\begin{aligned} \frac{\partial}{\partial t} \langle N_k \rangle &= \frac{2\pi}{\hbar} \sum_{k_1, k_2} M^2 \left[\langle N_{k_1} \rangle \langle N_{k_2} \rangle (1 \pm \langle N_k \rangle) (1 \pm \langle N_{k'} \rangle) \right. \\ &\quad \left. - \langle N_k \rangle \langle N_{k'} \rangle (1 \pm \langle N_{k_1} \rangle) (1 \pm \langle N_{k_2} \rangle) \right] \times \delta(E_{k_1} + E_{k_2} - E_{k'} - E_k), \end{aligned} \quad (7)$$

where the first product in the square brackets gives the “inflow” to state k , and the second term gives the “outflow,” and states k_1, k_2 , and k_3 label all other single-quantum states in the system. The δ -function gives energy conservation between the initial and final states involved in a collision. The term M is a number with units of energy that depends on microscopic details of the system; k' is not summed over because it is fixed by phase matching of the waves (momentum conservation, in particle language) of the three other states k_1, k_2 , and k .

In the limit of low density, when $\langle N_k \rangle$ is much less than unity for all k , this becomes

$$\begin{aligned} \frac{\partial}{\partial t} \langle N_k \rangle &= \frac{2\pi}{\hbar} \sum_{k_1, k_2} M^2 \left[\langle N_{k_1} \rangle \langle N_{k_2} \rangle - \langle N_k \rangle \langle N_{k'} \rangle \right] \\ &\quad \times \delta(E_{k_1} + E_{k_2} - E_{k_3} - E_k), \end{aligned} \quad (8)$$

which is the same as the classical, statistical equation written down by Boltzmann. In this form, it has the natural interpretation as a sum of probabilities of collisions of classical particles, in which the probability of a collision of particles with momenta k_1 and k_2 is proportional to the number of particles with each of those velocities. The principle of detailed balance says that in equilibrium, the inflow and outflow processes equal each other for every k . Solving for N_k under this condition then gives the equilibrium Maxwell-Boltzmann distribution of classical particles, in which $N_k \propto e^{-E_k/k_B T}$; in the full time-dependent solution of the quantum Boltzmann equation, the Maxwell-Boltzmann distribution corresponds to the $t \rightarrow \infty$ limit. Figure 1 shows a typical solution of the quantum Boltzmann equation for evolution from nonequilibrium to equilibrium. When the density is higher, so that $\langle N_k \rangle$ is not far below unity for all k , quantum statistics terms matter, so that the $(1 \pm \langle N_k \rangle)$ factors

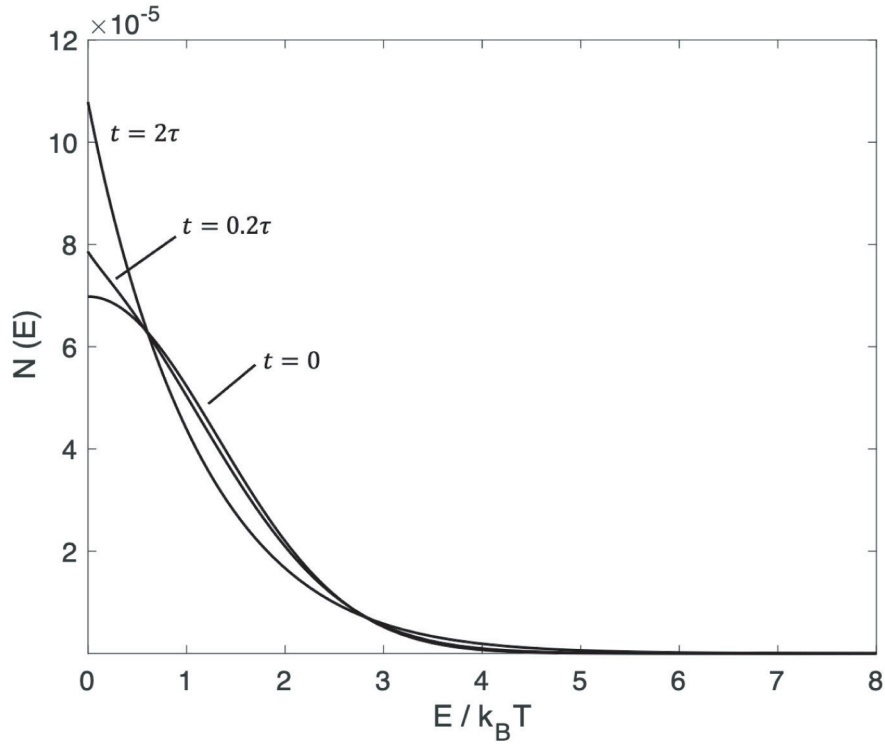


Figure 1. The solution of the quantum Boltzmann equation for various times after starting in a nonequilibrium initial state, for a low-density, two dimensional gas with collisional interactions. The times are given in terms of τ , the average collision time. Image credit: Hassan Alnatah.

must be included. In this case the principle of detailed balance in equilibrium gives the Fermi-Dirac and Bose-Einstein distributions for fermions and bosons, respectively.

As noted above, Ref. [5] showed that the quantum Boltzmann equation follows from the assumption that the off-diagonal terms of the density matrix are negligible. Going further, the time evolution of these off-diagonal terms was also calculated in Ref. [5], and it was shown that in most cases these terms decay rapidly to zero, consistent with the assumption used in the derivation of the quantum Boltzmann equation. Exceptions include “integrable” motion [6]; quantum “scars,” in which certain degrees of freedom can have periodic behavior [7]; and superconductors and superfluids, which can have spontaneous, long-lasting coherence (see, e.g., Ref., [4], Chapter 21). These exceptions can have long-term oscillatory behavior that appears time-reversible, but they will still have an overall increase of entropy subject to the second law, because they are always coupled, at least weakly, to the outside world. In particular, superfluids and superconductors have some range of states with very slow decoherence, to the degree that equations can be written for them that have no friction or damping terms. However, such systems are always coupled to other ranges of states with fast decoherence, and so they will eventually succumb to damping, however weak it may be. In practice, we can dismiss

the possibility of such exceptions in living systems, because they either require fine tuning to set up a “rigged” physical system, or they occur in nature only at very low temperature (where decoherence can become very slow), or both.

The quantum Boltzmann equation then implies the “H-theorem,” which is the basis of the second law of thermodynamics, namely, that entropy never decreases in a closed system. Technically, for a quantum system, the total entropy is given by the von Neumann entropy, which never changes. This definition is not too useful, however, and so other definitions of “effective” entropy can be used. The “semiclassical” entropy (also known as “diagonal” entropy [8]) is calculated using just the diagonal terms of the density matrix; for quantum particles at low density, this is

$$S = -k_B \sum_k \langle N_k \rangle \ln \langle N_k \rangle, \quad (9)$$

which is analogous to Equation (4) in Ref. [1]. At high density in quantum mechanics, this is modified to [9]

$$S = -k_B \sum_k \left(\langle \hat{N}_k \rangle \ln \langle N_k \rangle \mp (1 \pm \langle N_k \rangle) \ln (1 \pm \langle N_k \rangle) \right), \quad (10)$$

where the upper sign is for bosons and the lower sign is for fermions. For simplicity, we consider here the low-density (classical) limit, in which we drop all the terms in this formula with $(1 + \langle N_k \rangle)$. (It was shown in Ref. [5] that the second law is still derived if these terms are retained.)

Assuming conservation of the total number of particles, the time derivative of (9) is

$$\frac{\partial S}{\partial t} = -k_B \sum_k \frac{\partial \langle N_k \rangle}{\partial t} \ln \langle N_k \rangle. \quad (11)$$

Using the quantum Boltzmann equation (8), we then have

$$\frac{\partial S}{\partial t} = -k_B \sum_{k, k_1, k_2} C \ln \langle N_k \rangle [\langle N_{k_1} \rangle \langle N_{k_2} \rangle - \langle N_k \rangle \langle N_{k'} \rangle]. \quad (12)$$

where C is a positive number that contains the system-specific details of the collision process. For any choice of the four states k_1, k_2, k , and k' , the total of all terms in the sum (12) involving these four states is

$$\begin{aligned} & (\ln \langle N_k \rangle + \ln \langle N_{k'} \rangle - \ln \langle N_{k_1} \rangle - \ln \langle N_{k_2} \rangle) [\langle N_{k_1} \rangle \langle N_{k_2} \rangle - \langle N_k \rangle \langle N_{k'} \rangle] \\ & = \ln \left(\frac{\langle N_k \rangle \langle N_{k'} \rangle}{\langle N_{k_1} \rangle \langle N_{k_2} \rangle} \right) [\langle N_{k_1} \rangle \langle N_{k_2} \rangle - \langle N_k \rangle \langle N_{k'} \rangle]. \end{aligned} \quad (13)$$

If the in-scattering term in the square brackets is larger than the out-scattering term, then the denominator of the logarithm is larger than the numerator, making the logarithm negative. Conversely, if the in-scattering is less than the out-scattering term, the term in the square brackets is negative. Since the whole sum consists of terms like this, the total sum is less than or equal to zero, and therefore $\partial S / \partial t > 0$. This is the standard form of the H-theorem. Although we have shown this for the specific case of collisional interactions, it is easy to show that other dynamical processes lead to quantum Boltzmann equations that also have this same general result. The H-theorem is therefore entirely general.

Random statistics from measurements, collapse, or anything else played no role in this derivation; it is entirely the result of deterministic time evolution of the wave function of a system with many degrees of freedom. As discussed above, the crucial aspect of the calculation is fast decoherence, which occurs whenever there are many coupled states of a system.

2.3 The principle of detailed entropy increase

The result (13) does not just give an increase of the total entropy for a whole system. It says that every dynamical process gives an increase of the entropy; one might call this

the principle of “detailed entropy increase.” If all of the states involved (in the case given in Eq. (13), k_1, k_2, k and k') are all inside an observed subsystem of interest, then the entropy of that subsystem will always increase due to that process. If one or more of the states in a process is not inside the subsystem, then that particular process could give a decrease of entropy to the states that are inside the subsystem. But as the size of a system grows, its surface-to-volume ratio decreases. Processes that allow transitions to or from the outside of a subsystem occur at its surface. Therefore, for a sufficiently large system, surface processes can always be made negligible compared to processes within the bulk volume, and therefore the rule for the bulk, namely the second law, will prevail.

As a practical example, this means that it will never occur that a gas in a bottle sorts itself into regions of hot and cold, while also emitting heat to the outside of the bottle. That might satisfy a total entropy budget increasing, but it would violate the principle of detailed entropy increase. The gas in the bottle is a sufficiently large system to have to obey the second law on its own. Sufficiently large here, again, turns out to be not all that large. For the gas in the bottle to be large enough for the second law to apply, the dimensions of the bottle must be large compared to the “mean-free-path” of the gas, defined in Section 3, which can be very small, of the order of 100 nanometers in a gas at room temperature and atmospheric pressure.

These results can also be generalized to systems in which exiting a subsystem does not correspond to passing through a physical surface. For example, it is common in chemistry and biochemistry to analyze the populations of different molecules which can turn into each other via chemical reactions obeying a *mass-action equation* derived from thermodynamics. In this case, the molecules occupy the same spatial volume, but different populations can be treated separately. The principle of detailed entropy increase will still apply if at any point in time, most of the molecules are not in the process of leaving their population to enter another. We can visualize this as a “parameter space,” as shown in Figure 2, in which a “surface” represents a range of parameter values that allow conversion to or from other molecular species, and the “volume” represents ranges of parameters that do not. In general, whenever there are two or more parameters, the surface-to-volume ratio can be quite small.

Numerical results also indicate that in a non-closed system, the second law is not simply thrown out the window. To the degree the system approximates a closed, ergodically connected system, to that degree its behavior will approximate the second law. Figure 3 shows the long-time limit of the time-evolution of a low-density, Maxwellian gas, equilibrating through a collisional process as in Figure 1, but with two

added terms to each term in (8) to account for inputs and outputs from an external system, namely

$$\frac{\partial}{\partial t} \langle N_k \rangle = \dots + G - \frac{\langle N_k \rangle}{\tau_{\text{esc}}}, \quad (14)$$

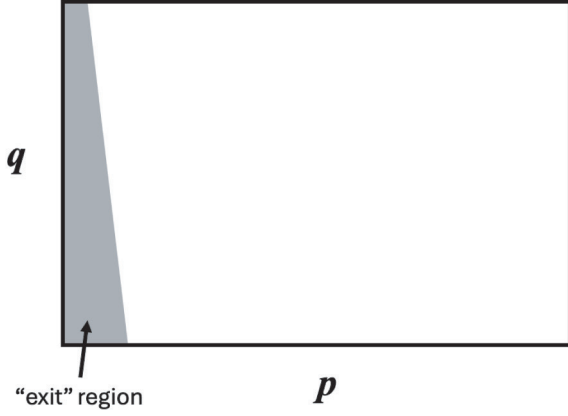


Figure 2. A generalized parameter space for a non-closed subsystem, with parameters p and q , in which only certain ranges of parameter values, which we can call “surfaces,” correspond to transition into or out of the subsystem.

where G is a fixed input rate from outside the system, and τ_{esc} is an intrinsic time scale for escape from the system. A system with these terms will reach a steady state when the total inputs and leaks balance. As seen in Figure 3, in the limit of steady state, when the escape time τ_{esc} is longer than the collision time τ , the distribution $N(E)$ is nearly equal to the equilibrium limit for a system with no decay; the distribution does not deviate much from the equilibrium shape until the time for escape τ_{esc} is much less than the collision time τ . The general principle extracted from many studies of nonequilibrium systems (e.g. Ref. [10]) is that a nearly-closed system with a slight coupling to the outside world has only a slight deviation from the behavior of a closed system.

2.4 The problem of irreversible behavior in a quantum system

The quantum Boltzmann equation and the associated H-theorem have been the subject of controversy for over 100 years (See., e.g., Ref. [11]). The basic problem is that quantum mechanics is fundamentally a time-reversible set of equations, while the second law gives irreversible behavior. The quantum Boltzmann equation will always give unchanging, steady-state behavior in the $t \rightarrow \infty$ limit.

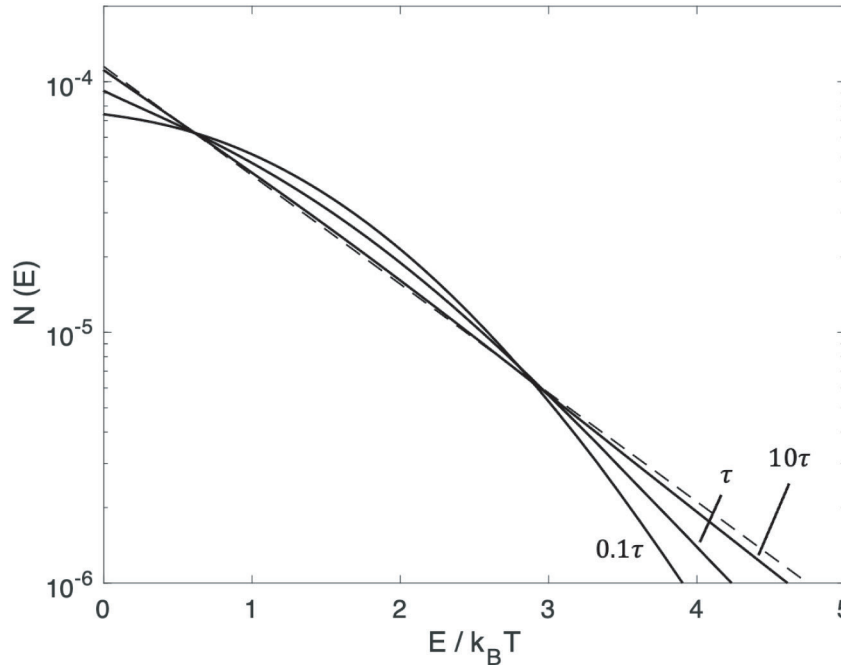


Figure 3. Log plot of the solution of the quantum Boltzmann equation for collisional equilibration in the $t \rightarrow \infty$ limit, for a low-density, two-dimensional gas, for three values of the leakage time to the outside of the system, offset by a steady-state replenishment with a $N(E)$ profile of the same form as the initial state of Figure 1, to keep the total density constant. The solid lines correspond to different values of the leakage time in units of the average collision-scattering time, τ . The dashed line corresponding to the infinite-lifetime case is equal to a thermal Maxwell-Boltzmann distribution, $N(E) \propto e^{-E/k_B T}$. Image credit: Hassan Alnatah.

We can summarize the state of modern thinking about the problem of irreversibility in the following points:

- For any *finite* system, the irreversibility of the Boltzmann equation cannot be strictly correct, because any finite system is subject to the Poincaré theorem, which says that all finite systems have a finite recurrence time to repeat all behaviors. (For the quantum mechanical version of the proof of this theorem, see Ref. [4], Section 18.1.)
- For an *infinite* system, the Poincaré recurrence time goes to infinity, and there is no recurrence. Since every local system here on earth system is actually coupled to the outer universe through radiation into outer space, even if very weakly through slow processes like quantum tunneling, then if the universe is truly infinite, then the second law is exact.
- For a large but not infinite system, the Poincaré recurrence time becomes extremely long, even for systems that don't seem so large; for example, for 100-1000 atoms in a gas, the Poincaré recurrence time can be much longer than the age of the universe.
- Although the quantum Boltzmann equation is only approximate for a finite system, it is a very good approximation for all realistic time scales. This is because in the derivation of the quantum Boltzmann equation, phase coherence terms are set to zero, and in the physical world, these terms are infinitesimally small.
- Ref. [5] showed self-consistency of the derivation of the quantum Boltzmann equation by calculating the time evolution of the most important phase coherence terms that arise in the derivation, and showed that these tend toward zero rapidly. A technical point is that the calculation of these coherence terms requires the assumption that other, higher-order coherence terms are negligible, but these higher-order terms were not calculated. Therefore the derivation of the quantum Boltzmann equation requires the unproven assumption that high-order correlation terms are small compared to lower-order terms, a very general procedure in physics known as a “perturbation expansion.” This is the same assumption that is used to derive the quantum field theory itself (see, e.g., Refs. [12], [13], or Chapter 15 of Ref. [4]). If this assumption is not true, then the whole structure of quantum field theory itself breaks down.
- The fact that physics is time reversible means that if we took the exact quantum state of the system immediately after the initial state, and time reversed it, we would see evolution back to the initial state. But the quantum Boltzmann equation will never give backwards-in-time evolution, for the same reason that it does not give Poincaré recurrence, because

the information in off-diagonal elements of the density matrix is set to zero, and this information is crucial for recovering the reversed-time behavior. This means that if we did have a way to keep and encode all that off-diagonal information, we could create an initial state that did not obey the quantum Boltzmann equation. But that would be the same as extreme fine-tuning of the initial state of the system to an incredibly high degree, keeping track of phase relationships of 10^{23} particles or more.

3 The second law in spatially inhomogeneous systems

3.1 Summary of thermodynamic flow results

The discussion of Section 2.3 did not explicitly discuss spatial variation, other than to introduce the idea of boundaries around subsystems. Much of our intuition about machines involves irreversible spatial flow, e.g., from hot to cold, from high pressure to low pressure, high density of some species to low density, etc. These behaviors can be derived directly from the second law (and therefore, implicitly, from the quantum Boltzmann equation). This is because the second law implies the *diffusion equation*,

$$\frac{\partial}{\partial t} n(\vec{x}, t) = D \nabla^2 n(\vec{x}, t), \quad (15)$$

where $n(\vec{x}, t)$ gives the density of particles (or, more generally, the energy density) at location $\vec{x}$ as a function of time, and D is the *diffusion constant*, which depends on the microscopic properties of the system. Figure 4 shows a typical solution of the diffusion equation. When long-range forces are allowed, the diffusion equation becomes the *drift-diffusion equation* (see Ref. [14], Section 5.8). Diffusion, drift, and the change of the energy distribution due to the quantum Boltzmann equation discussed in Section 2.3, are all accounted for in the general *Boltzmann transport equation* (See Ref. [14], Section 5.11).

As seen in Figure 4, the diffusion equation gives irreversible behavior in time, as the spatial profile of a gas evolves toward equilibrium, which for the spatial distribution means that $n(\vec{x}, \infty)$ equals a constant. When $n(\vec{x}, \infty)$ represents a density of particles, we equate this with the *chemical potential* being constant everywhere in space, while if $n(\vec{x}, \infty)$ represents an energy density, we can equate it with the temperature being constant.

The derivation of the diffusion equation is given in Section 3.2. This equation, and the more general drift-diffusion equation, hold true when the timescales for flow are long

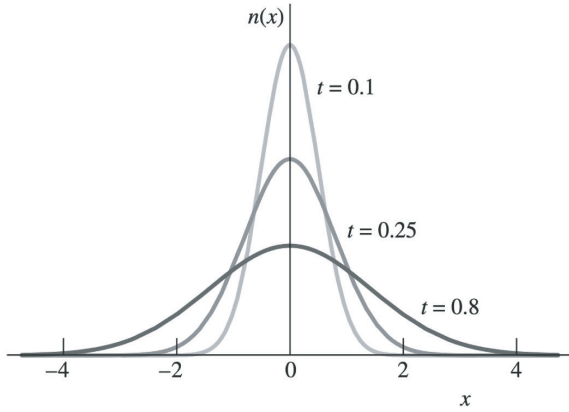


Figure 4. The solution of the diffusion equation (15) for a nonequilibrium initial state, for $D = 1$. From Ref. [14].

compared to the *relaxation time*, or thermalization time, which is computed via the quantum Boltzmann equation; this is the same timescale as the *decoherence time* for decay of off-diagonal terms of the density matrix, discussed in Section 2.2. Thus, the diffusion equation and its cousins give irreversible behavior in spatial flow under the same conditions and for the same underlying reason as the quantum Boltzmann equation, and can be seen as the spatial version of the second law of thermodynamics. In common terms, it implies that heat flows from hot to cold and not vice versa, and particles flow from high concentration to low concentration, and not vice versa. This type of flow will stop when all parts of the system are at the same temperature and chemical potential.

The concept of “friction” comes from the same microscopic quantum description used in the diffusion equation. Friction also comes ultimately from the same decoherence processes as accounted for in the quantum Boltzmann equation, but describes an effective opposing force, also called “drag,” when there is a driving force on a system. The *Einstein relation* and the *fluctuation-dissipation theorem* (see, e.g., Ref. [14], Sections 5.8 and 9.8) both express the same principle, namely that friction opposing a driving force, and diffusive motion due to random fluctuations in the absence of a force, both arise from the same underlying dissipation, with the same characteristic relaxation time, given by the quantum Boltzmann equation. For a standard drag force, if an object is set into motion and then the driving force is removed, the velocity of the object given by

$$\frac{dv}{dt} = -\frac{1}{\tau}v, \quad (16)$$

where v is the velocity and τ is the relaxation time. The equation has the solution $v(t) = v(0)e^{-t/\tau}$, which means that

a velocity will decrease irreversibly by a factor of $1/e$ in an amount of time τ .

3.2 Derivation of the diffusion equation using the quantum Boltzmann equation

A derivation of the diffusion equation in terms of microscopic particles is given in Ref. [14]. Here, it is shown that the concept of compact particles is not crucial; the derivation can be done in terms of quantum wave equations, just as was done for the quantum Boltzmann equation.

We start with the relaxation time τ , which is computed using the Boltzmann equation; this is the time for a single state to be depleted by collisions or other thermalizing interactions. This time constant depends on the details of the specific system, and in general also has different values for different subsets of a system, e.g., it can be much longer for low-momentum states than for high-momentum states. In equilibrium, this time constant is simply computed from the sum of the out-scattering terms in (8), that is, the second, negative term in the square brackets.

We can then define the *mean free path* l as the velocity of a given quantum state times the relaxation time, $l = v\tau$. (Per the discussion at the outset of this paper and presented in Ref. [4], it is convenient to define the velocity used here as the velocity of a particle, but not necessary; one can simply define it in terms of the wavelength of the appropriate wave function, that is, $\vec{v} = \vec{p}/m = -i(\hbar/m)\nabla$.) The mean free path is the average distance a quantum particle travels *ballistically*, that is, without much randomization of its motion; in wave terms, this is the same as the wave *coherence length*, that is, the distance over which the wave has a well-defined phase.

Spatial inhomogeneity was not explicitly addressed in Section 2.2, although the approach is quite general. We can allow for spatial inhomogeneity explicitly by writing, for any given point in space x , the contribution of waves with momentum k coming from a distance x :

$$\frac{\partial N_k}{\partial t} \propto e^{-x/lk} = e^{-x/v_k\tau} = e^{-xl(\hbar k/m)\tau}. \quad (17)$$

This reflects the fact that particles will scatter out of their ballistic motion with a rate of $1/\tau$, the relaxation time given by the Boltzmann equation. A constant rate of decay gives an exponential decay, with the contribution of the system suppressed from distances further away.

If we imagine a surface somewhere in a gas, defined as $x = 0$, the total input of momentum is then the sum of the positive

momentum from the left and the negative momentum from the right,

$$\sum_{k=0}^{\infty} (\hbar k) \int_{-\infty}^0 dx' N_k(x') e^{x'/l_k} + \sum_{k=0}^{\infty} (-\hbar k) \int_0^{\infty} dx' N_k(x') e^{-x'/l_k}. \quad (18)$$

We assume that the gas is in a thermal Maxwellian distribution, with a density that is slowly varying compared to the mean free path, so that we can write

$$N_k(x) = N(x) e^{-E_k/k_B T}, \quad (19)$$

with

$$N(x) = N_0 + \left. \frac{\partial N}{\partial x} \right|_{x=0} x. \quad (20)$$

Writing $N_k(x)$, with both k -dependence and x -dependence, implicitly assumes a quantum formulation like a Wigner distribution, in which the momentum-space k -distribution is defined by Fourier analysis of a localized (but not infinitely small) spatial region.

Adding the two terms of (18) together, and changing variables in the first integral from $x' \rightarrow -x'$, the leading-order terms cancel out, and we are left with

$$\begin{aligned} & -2 \left. \frac{\partial N}{\partial x} \right|_{x=0} \sum_{k=0}^{\infty} (\hbar k) \int_0^{\infty} dx' x' e^{-x'/l_k} e^{-E_k/k_B T} \\ & = -2 \left. \frac{\partial N}{\partial x} \right|_{x=0} \sum_{k=0}^{\infty} (\hbar k)^2 l_k e^{-E_k/k_B T}. \end{aligned} \quad (21)$$

To get the average momentum, we divide this by the total density flowing in from each side, which is to leading order,

$$N(0) \sum_{k=0}^{\infty} \int_{-\infty}^{\infty} dx' e^{-E_k/k_B T} e^{-|x'|/(l_k/m)} = 2N(0) \sum_{k=0}^{\infty} e^{-E_k/k_B T} l_k. \quad (22)$$

Performing the integrals over k , with $E_k = (\hbar k)^2/2m$, then gives the average velocity

$$\langle v \rangle = \frac{1}{m} \langle p \rangle = -\frac{1}{N(0)} \left. \frac{\partial N}{\partial x} \right|_{x=0} k_B T \frac{\tau}{m}. \quad (23)$$

This is sometimes called “Fick’s law”—the average momentum is proportional to and opposite the density gradient, times the diffusion constant $D = k_B T \tau$. It can be written in vector form as

$$n \langle \vec{v} \rangle = -D \nabla n, \quad (24)$$

where n is the density per volume.

The diffusion equation is then derived by combining this with the *continuity equation*

$$\frac{\partial n}{\partial t} = -\nabla \cdot n \langle \vec{v} \rangle, \quad (25)$$

which says that spatial divergence of the flow gives the change of density at any point. Using (24) for $\vec{v}$ gives us

$$\frac{\partial n}{\partial t} = D \nabla^2 n, \quad (26)$$

which is the diffusion equation.

Although we have derived this here for particles with mass, the same approach can be used to derive a diffusion equation for *phonons*, that is, the quanta of vibration that make up heat. There is therefore the same type of diffusion equation for heat flow.

4 The role of gates in counter-entropic processes

4.1 What is the essential characteristic of life and machines?

With the considerations of the previous sections in mind, we can now address a very basic question: what is the essential characteristic of a living system, and by extension, a machine made by something living, or some other artifact of a living system?

Schrödinger, in his book *What is Life?* [15], talked of life as an essentially nonequilibrium process. But it is not adequate to simply equate life with nonequilibrium. The planets orbiting the sun, for example, are not in equilibrium, as there is a characteristic time for the solar system to expel planets [16]. But we would not say that they are a living system. Others have talked of life as “homeostasis” [17], that is, continuity of existence of a species, but simply existing without change is not unique to living systems; after all, an igneous rock is also stable, but is not living. Other definitions (see, e.g., Ref. [18]) often involve details specific to known living systems, such as reproduction or carbon.

We can instead use a modified version of Schrödinger’s definition and say that a machine system, and a living system, is one in which a *subsystem violates the second law of thermodynamics over time scales long compared to the relaxation time, even though the whole system does not*. In

other words, it is one in which heat flows from cold to hot (a refrigerator), or concentration flows from low to high (a pump), or motion changes from random to net linear or circular collective motion (an engine, or motor), and this is not a mere fluctuation, but a stable behavior.

In Section 2.3, we discussed the principle of detailed entropy increase, that not only whole systems, but also subsystems, obey the second law. How then can we say that machines exist in which subsystems violate the second law? The key is found the summary statement given of Section 2.3, “The entropy not only of a whole system, but of every sufficiently large, *ergodically connected* subset of a system always obeys the second law of thermodynamics.” A key assumption used in the derivation of the H-theorem was that the system under consideration was ergodically connected—that is, interaction processes mix all of the states of the subsystem freely. Therefore, for a subsystem to violate the second law, it must have domains that are not ergodically connected, that are walled off, so to speak.

For shorthand, we can call this *counter-entropic* behavior, meaning that in some subsystems, entropy flows oppositely to what it would do if the subsystem were fully ergodically connected. Designed machines and living systems have elements that push back against normal entropic flow.

Let us look in detail at the operation of an engine following a Carnot cycle, which is the most efficient possible machine [19]. We imagine a cylinder with a piston that can freely move up and down, with two external heat reservoirs, as shown in Figure 5. The piston has a constant downward force applied. In stage 1 the beginning of the cycle, this cylinder is given a heat link to the hot region. Heat flows in from this reservoir, raising its pressure and causing the gas in the cylinder to expand. This pushes the piston down. In stage 2, a switch turns off the thermal connection to the hot reservoir. The gas continues to push on the piston, but now, without any heat input, the temperature of the gas falls as it expands. In stage 3, a heat link is opened to the cold region. Heat flows out of the gas to cold region, which causes its pressure to drop, which then leads to the piston pressing back on it, due to the external force on the piston. In stage 4, the connection to the cold reservoir is switched off. The piston continues to press on the gas to compress it, but now the temperature of the gas rises due to the compression.

At every stage of this process, heat always flowed from hot to cold; to put it another way, each ergodically connected subsystem at each moment in time obeyed the second law. However, if we restrict our attention to the subsystem comprised of the gas, piston, and hot reservoir alone, and

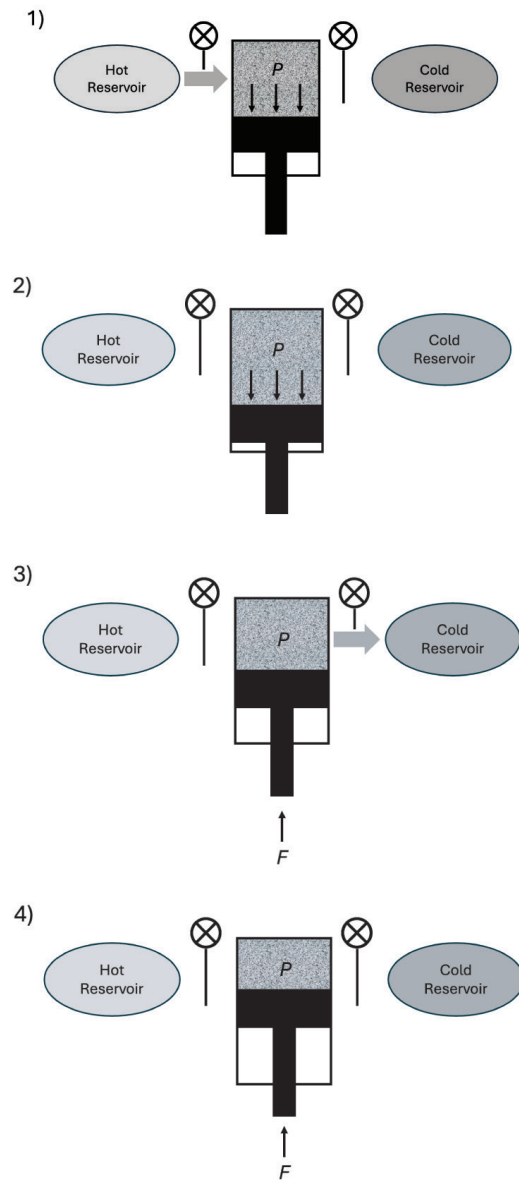


Figure 5. Standard model of the four cycles of a Carnot engine: 1) isothermal expansion with heat input, 2) adiabatic expansion with temperature drop, 3) isothermal compression with heat input, and 4) adiabatic compression with temperature increase.

exclude the cold reservoir, the second law is violated for that subsystem. Random motion in the hot reservoir (microscopic heat motion of atoms) has been converted into ordered linear motion. This does not violate the theorem of Section 2.3, because the hot reservoir and the piston are not continuously ergodically connected.

But notice the crucial role of “switches” in this process. The cycle only works because the engine is positioned

on an interface between two heat reservoirs, and the thermodynamic flow between the machine and the reservoirs can be turned on and off. At least two switches with binary action are needed, to turn on and off the flow between the device and the two reservoirs on each side. Implicit in this is that two “detectors” with “memory” (a.k.a. “information”) are needed, to turn on and off the switches at the right points in the cycle, based on response to the state of the system, and keep it switched that way until the next switch is triggered.

As shown in Ref. [1], this process has the same efficiency as a Szilard engine [20], in which Maxwell’s demon converts energy flow between two heat reservoirs into usable linear motion. This is not accidental, because a Carnot engine can be viewed as a type of Maxwell’s demon with two bits of information storage. Each switch detects some information about the state of the system, namely which stage of the cycle it is in, and produces a macroscopic response (heat flow or no heat flow). In the same way, the Szilard engine relies on a switchable interface: a door is opened to allow flow between two reservoirs at different pressures. It detects two aspects of the system (incoming atom on the left, or incoming atom on the right) and moves a door in response.

In some actual machines, it is not always so easy to see where these “switches” reside, but something always plays that role. For example, in an internal combustion engine, heat input is switched on and off by the timing of the spark plug and injection of fuel, which gives an explosion of heat only a specific times. Instead of heat flow out being switched on, another switch allows thermodynamic flow of high pressure gas to a low pressure region (exhaust); heat flows continuously out during the whole cycle through the “block” of the engine. The same is true of kitchen refrigerators. In this case the gas is cycled around a loop, and some regions of the pipe in the loop have strong coupling to a hot region (the grille on the back of the refrigerator, which radiates heat), and some regions have strong coupling to a cold region (passing along the inside of the refrigerator). One can call these methods of turning on and off thermodynamic flow “clever design,” as they prevent to a large degree any unwanted direct flow between the two reservoirs.

I assert that *every* machine has this feature, of having something that switches thermodynamic flow between two reservoirs. These must include 1) an *interface* between two reservoirs that prevents mixing, and 2) a “gate” that can turn on and off the flow between these reservoirs, with a high degree of change of the flow rate. The switch does not need to be a perfect switch—there may be some residual thermodynamic connection between the the regions—but if it is not, the

efficiency of the machine will drop from the optimal value obtained by following a Carnot cycle, which requires perfect switches. Effectively, a machine with imperfect switches will act the same as a perfect Carnot machine plus a direct heat leak from the hot to the cold reservoir.

We have focused so far on thermodynamic flow of heat, but the same behavior of switching leading to apparent violation of the second law can be accomplished by other types of thermodynamic flow; for example, by number-flow from regions of high concentration (technically, high “chemical potential”) to low concentration. We may therefore extend this discussion to biochemical systems that control the rate of flow from one “population” to another. Although two types of molecules may intermingle in the same space, they may effectively be uncoupled reservoirs as long as the chemical reaction that leads to conversion between them can be turned off. A “switch” may then be the presence of a third molecule that rapidly increases or decreases the rate of conversion of molecules from one type to the other.

It may be objected at this point that spontaneous emergence contradicts my assertions. The natural world has various examples in which some subsystem violates the second law in the same way as described above [21]. For example, in the natural water cycle of the earth, water in vapor form, with random motion, is converted to having linear motion in the form of rain fall and river flow downhill. The reason why this happens is the same as described above—physical motion from one place to another changes what heat reservoirs the water is ergodically connected to. The hot surface of the earth causes the vapor to rise, which then disconnects it from the hot surface of the earth and connects it to the cold upper atmosphere, cooled ultimately by the cold vacuum of outer space. Note that the two aspects discussed above are present: an interface (the surface of the earth) with suppressed thermal connection between its two sides, and a means of switching on and off the thermal connection to the two sides of the interface (gravity force that causes hot vapor to rise and rain to fall), on a time scale short compared to the time scale for thermal flow between the two reservoirs.

Instead of trying to exclude this type of example from our definition of machines, we can instead simply agree to call this an inefficient simple machine. As discussed in Ref. [1], this sort of machine can arise spontaneously because there is a natural instability in the system, which in this case is the inversion of a cold region above a hot region in a gravity field. This is a system of low entropy compared to a homogeneous system, if the entropy of the expelled mass to form the planet is excluded. An entropy analysis then says that the probability

of random motion of water vapor turning into linear motion is greater than the probability of staying in the initial low-entropy state of all the heat on one side of the interface, on the ground. But this system has exhausted all of the available resources for spontaneous machinery. There is a natural length scale given by the size of a convection cell in the temperature-inverted system, and after convection appears, there are no more natural length scales to exploit.

We can instead quantify the *degree of machinery* in a system by the number of types of counter-entropic processes it contains (many copies of the same process do not count, e.g. many convection cells acting the same way). A living system, or an engineered system, is one in which the number of different processes is much greater than 1. While I am not aware of any attempt yet to count this for actual living systems, the number clearly must be at least in the hundreds or thousands.

Note that the existence of gates is a *necessary* condition for counter-entropic behavior by the above argument, but not *sufficient*. In addition to the gates themselves, all of the above examples include *detectors* that control the opening and shutting of the gates based on detected information about some conditions of the system, and systems of *routing* that direct signals from detectors to the proper gates. These detection and signaling systems may be crucial in creating higher degrees of machinery. The characteristics of detectors and their signaling to gates is beyond the scope of this present work.

4.2 The essential features of thermodynamic gates

As we have seen, gates, or thermodynamic switches, are essential for counter-entropic behavior, to connect disconnect thermodynamic flow. Figure 6 illustrates the general operation of a gate. There are several essential features of every gate. First, there must be an *interface* with a high barrier between what would otherwise be rapid thermodynamic flow accompanied by the increase of entropy. As discussed in the previous section, this can be flow from a higher concentration of thermal energy to lower concentration, or higher concentration of particles to lower concentration, or from one species of molecule to another via a chemical reaction from high chemical potential to lower chemical potential.

Second, there must be a controllable portal, or *gate*, that can change the rate of flow, or diffusion, across the interface. From the laws of thermodynamics, we know that a Carnot cycle has the highest possible efficiency [19], and a Carnot cycle corresponds to abrupt switching off of connections to the reservoirs. In the case of the piston engine described in Section 4.1, this corresponds to abrupt switching between stages with no energy input or output (“adiabatic” change)

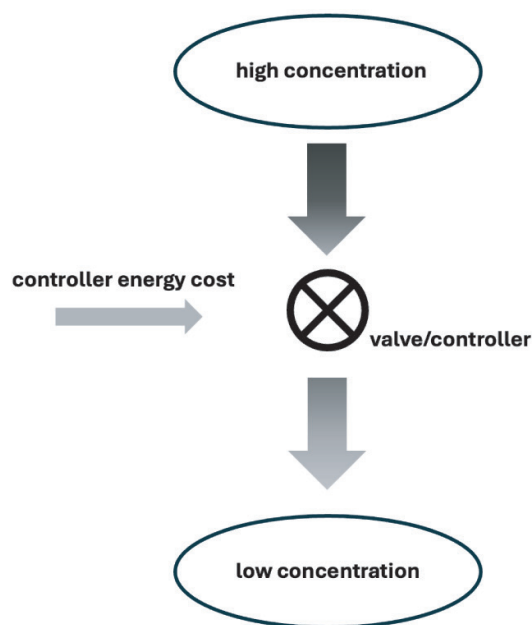


Figure 6. Illustration of the action of a general gate across a thermodynamic interface.

and stages with change of volume at constant temperature (“isothermal” change). If the switching is not abrupt, this will correspond to an additional flow from one top reservoir to the other without any effect on the machine other than wasting free energy. We can therefore characterize a switch as having a time scale for operation that is short compared to the time scale for flow between the reservoirs that bypasses the gate.

A third feature is that a good gate has a *low energy cost* to operating the controller. There is always some energy dissipated by the controller, for the same reason as the Landauer argument, namely that resetting the state of the controller is an irreversible process (see Ref. [1]). As discussed in Ref. [1], for the Szilard engine using Maxwell’s demon, the cost of opening and shutting the valve itself is typically assumed to be negligible. The only strictly unavoidable energy cost is the cost of “memory” of the state that is set.

We thus see that high efficiency corresponds to information in “bits” with sharply distinguished states, which we can call 1 for the “on” state and 0 for the “off” state of the controller. In the case of the Carnot engine, we needed two bits of information to keep in memory which of the four stages of the cycle the machine was in.

We can therefore talk of an “on/off ratio” that characterizes the efficiency of any switch, or gate. Flow across the interface

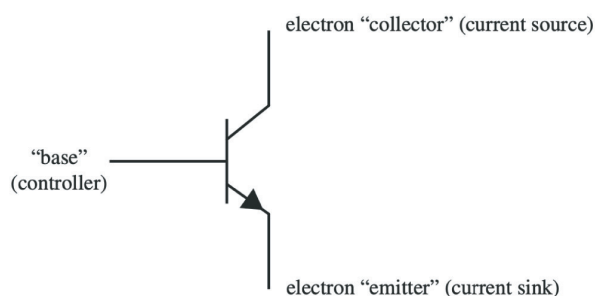


Figure 7. Typical transistor symbolism and terminology. From Ref. [23].

in the “off” state corresponds to inefficiency, and therefore optimality implies the highest possible on/off ratio. The assumption of optimality, that is, good design, promoted by Bill Bialek [22] and others, which has been quite productive experimentally, leads one to expect that biological switching will have high on/off ratio.

It may be obvious to some readers at this point that the switches we have been discussing have the same behavior as electrical transistors. Figure 7 shows a standard diagram of an electrical transistor, which is simply an electronic switch. (For extended discussion of transistor electronics, see Ref. [23]), Chapters 4 and 5.) The input “gate,” also sometimes called the “base,” controls the current between the “source” and “drain,” also called the “emitter” and “collector” in some devices. (An oddity of electronics is that current flow is defined oppositely to electron flow, for historical reasons.) The current flow between the source and drain is fundamentally thermodynamically driven, from high electron chemical potential to low electron chemical potential, just as described in Section 4.2 for a generic thermodynamic gate.

Transistors can come in two varieties: default-on, and default-off. In other words, the gate can be left normally open to thermodynamic flow, and energy applied to the controller to shut it off, or the gate can be left normally closed, and energy applied to the controller to open up the flow. Both of these options have ubiquitous use to implement conditional logic. The same behaviors occur in living systems, as discussed in Section 4.3.

We can also distinguish between *active* switches and *passive* switches, which may also be called feedback and feed-forward switches. Active switches have some type of mechanism to detect the present state of the system and control the gate in response, for example, a thermostat controlling the on-switch of a heater. Passive switches, on the other hand, control the flow across an interface using deterministic routing of what

will happen in a system. For example, in a refrigerator, gas is cycled around a loop, and hits the exterior heat-conduction fins on the back to dump heat just at the right time in its cycle. This latter, passive type is effective when the behavior of a system is very predictable, and so can be built into the design in advance; the former, active type of switching is needed when the system must respond to an unpredictable variable.

A characteristic of many efficient switching systems is that the output of one switch can become the active controller of one or more other switches. This allows extensive *networks* to be built up of connected machines. These are seen in the complicated “circuit diagrams” of modern electronics. Long before electronics came along, the same type of network signal- and-switch logic was implemented in mechanical systems such as clocks and adding machines.

4.3 Gates in living systems

We have seen that the interface+gate structure is essential for counter-entropic behavior. When we look at biological systems, we indeed see many examples of this type of structure. The first and most obvious is the physical structure of living things: *organisms* with well-defined surfaces or interfaces defining an inside and outside of the organism; e.g., the skin. These surfaces have interfaces that allow switched flow from outside to inside (e.g., ingestion) and from inside to outside (e.g., waste removal). Inside each organism we see the same structure: *organs* with gated membranes; inside these, *cells* with cell walls and portals (the terms “portal” and “pore” come from the Latin word for a gate); and *organelles* inside cells. At levels above the organism, we often see *colonies* of organisms with well-defined boundaries.

As mentioned at the end of Section 4.2, in human-designed electronics and machinery, it is common not only to have few-gate simple machines, but networks of switches that implement many options for responses. The same occurs in biological systems. Many biologists are familiar with the complicated “biochemical pathways” diagrams of living systems (see, e.g., the site given in [26]). In this case, the subsystems are not usually spatially localized behind physical surfaces, but are populations of things (which can be as small as simple molecules, or large proteins, or even whole machines) that are stable against conversion into other species unless a “gate” process is turned on that allows quick conversion, i.e., exit from the population via an exothermic (thermodynamically “downhill”) chemical reaction. The gate in this case does not sit at a physical interface, but acts as the trigger for allowing or disallowing conversion between two populations, which are effectively two thermodynamic reservoirs.

We now begin to see that the complicated chemical pathway diagrams of living systems are not just *analogous* to the switch-and-signal diagrams of electrical circuits (as argued, e.g., in Ref. [27]); they really *are* switch-and-signal diagrams. This network behavior can be so complicated that we lose sight of the fact that each switch-and-signal operation is still essentially *gated thermodynamic flow*. Any system that has irreversibility is governed by the second law of thermodynamics. In a switch of any type, irreversibility occurs in two ways. One is the closing or opening of a switch without bouncing open again; this is equivalent to the required dissipation of energy of the “memory” element, in the Landauer/Maxwell-demon argument. The second time-asymmetric process is the flow in only one direction, from one reservoir to another, when the gate is open.

To get a high on/off ratio, the generic mechanism is to have *triggered bistability* of the rate of thermodynamic flow [28]. This is often implemented via *positive feedback*, in which a thermodynamic flow (e.g., a chemical reaction) causes something to occur that accelerates the flow. Positive feedback, and bistability with high on/off ratio between two states, is well documented in many biological systems (for a review, see, e.g., Ref. [29]).

It is beyond the scope of this paper to review how switches in all biological systems operate and where they are located. We can instead simply make an observation and a conjecture. The observation is that, as shown in the previous sections, switching behavior is *necessary* for the counter-entropic behavior that is ubiquitous in life. Therefore, where we see such counter-entropic behavior, we should *look for* switches, both active and passive, operating on interfaces or controlling flows between molecular populations. These counter-entropic flows include *pumps* (such as a hydrogen pumps that give flow against a concentration gradient), *engines* (collective linear or circular motion generated out of random motion), and in some cases, *refrigerators* (heat flow from a colder region to a hotter environment, to resist overheating). In many real biological systems, it is already known that these actions involve fine tuning to control channels of entropic flow, e.g., the complex design of lipid membranes in eukaryotes [24, 25]. The thermodynamic considerations presented in this paper imply that we should expect well-controlled switching behavior to be universal.

Second, we can make the *conjecture* that these switches will be highly efficient, that is, have high on/off ratio and require little energy to operate, based on the general principle of optimization of Bill Bialek mentioned above, and the knowledge that the efficiency of counter-entropic processes requires efficient switches, as discussed in Section 4.1 in regard to the Carnot engine.

5 Conclusions

Putting together the conclusions of Ref. [1] and this paper, we can make the following statements:

- The second law of thermodynamics applies not only to systems as a whole, but also to subsystems. We cannot simply wave a wand and say that if the total entropy of a system increases in a whole system, there is no surprise if some subsystem has a dramatic decrease of entropy.
- Living systems and machines designed by humans have subsystems that appear to violate the second law of thermodynamics. They do this by a specific boundary+triggered-gate process that allows thermodynamic flow only in certain directions at certain times.
- For the second law to hold true in such a subsystem, it must be the case that the existence of the specific boundary+triggered-gate process itself is a low-entropy state.
- Spontaneous emergence of counter-entropic machines at an interface can occur, but only at the simple degree of what is allowed by a natural instability of that interface, such as the gravity well of a planet that creates a separation between a high temperature planet surface and low-temperature outer space, leading to convection cells. Thus, spontaneously emerging machines are limited to very simple behavior.
- All the higher-degree counter-entropic systems we know of are generated by prior existing living or machine-like systems using a process that itself is counter-entropic (e.g., human design of machines, or reproduction of offspring by living organisms). This generation process therefore presumes the existence of a prior initial state with even lower entropy. This presents a serious challenge for physical models of the origin of life, since to form spontaneously, each separate gated-interface process would need its own, independent natural instability, and the probabilities of each occurring must be multiplied.

This analysis in terms of counter-entropic flow and thermodynamic gates has utility in how we understand living systems. We can generally say the following:

- Where we see counter-entropic behavior, we should expect to see gate/switch behavior.
- Where we see gate/switch bistable behavior, we should expect to see counter-entropic behavior.
- Since efficient switching is needed for a Carnot cycle, which is the most efficient possible machine cycle, the principle of optimality in biology implies that we should expect efficient switching with high on/off ratio in biological systems.

- Since the switching behavior of biological networks is of the same fundamental nature as that of human-designed electrical and mechanical switching networks, we should expect that systems engineering methods will work well for biological networks. This project is already well under way [30, 31, 32].

References

- [1] D.W. Snoke, "Spontaneous appearance of life and the second law of thermodynamics," *Biocosmos* **2**, 1 (2022).
- [2] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development* **5**, 183 (1961).
- [3] G. Sewell, "On 'compensating' entropy decreases," *Physics Essays* **30**, 1 (2017).
- [4] D.W. Snoke, *Interpreting Quantum Mechanics: Modern Foundations*, (Cambridge University Press, 2024).
- [5] D.W. Snoke, G.-Q. Liu, and S.M. Girvin, "The basis of the second law of thermodynamics in quantum field theory," *Annals of Physics* **327**, 1825 (2012).
- [6] A.L. Retore, "Introduction to classical and quantum integrability," *Journal of Physics A: Mathematical and Theoretical* **55**, 173001 (2022).
- [7] M. Serbyn, D.A. Abanin, and Z. Papić, "Quantum many-body scars and weak breaking of ergodicity," *Nature Physics* **17**, 675 (2021).
- [8] A. Polkovnikov, "Microscopic diagonal entropy and its connection to basic thermodynamic relations," *Annals of Physics* **326**, 486 (2011).
- [9] W. Band, *An Introduction to Quantum Statistics*, (D. van Nostrand, 1955).
- [10] J. Keeling, "Superfluid density of an open dissipative condensate," *Physical Review Letters* **107**, 080402 (2011).
- [11] H. Brown, "Once and for all: the curious role of probability in the Past Hypothesis," *PhilSci archive preprint* 13008.
- [12] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems*, (McGraw- Hill, 1971).
- [13] F. Mandl and G. Shaw, *Quantum Field Theory*, 2nd ed., (Wiley, 2010).
- [14] D.W. Snoke, *Solid State Physics: Essential Concepts*, 2nd ed., (Cambridge University Press, 2020).
- [15] E. Schrödinger, *What Is Life? The Physical Aspect of the Living Cell*, (Doubleday, 1944).
- [16] Y. Lithwick and Y. Wu, "Secular chaos and its application to Mercury, hot Jupiters, and the organization of planetary systems," *Proceedings of the National Academy of Sciences (USA)* **111**, 12610 (2014).
- [17] A. Diéguez, "Life as a homeostatic property cluster," *Biological Theory* **7**, 180 (2013).
- [18] C. Mariscal, "Life," *Stanford Encyclopedia of Philosophy* (Winter 2021).
- [19] F. Reif, *Fundamentals of Statistical and Thermal Physics*, (McGraw-Hill, 1965).
- [20] L. Szilard, "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings," *Zeitschrift für Physik* **53**, 840 (1929).
- [21] "Emergence in solid state physics and biology," *Foundations of Physics* **50**, 1098 (2020).
- [22] W. Bialek, "Ambitions for theory in the physics of life," *Arxiv*:2401.15538.
- [23] D.W. Snoke, *Electronics: A Physical Approach*, (Pearson, 2015).
- [24] T. Dingjan and A. Futerman, "The fine-tuning of cell membrane lipid bilayers accentuates their compositional complexity," *BioEssays* **43**, 2100021 (2021).
- [25] J. H. Lorent, K. R. Levental, L. Ganesan, G. Rivera-Longsworth, E. Sezgin, M. Doktorova, E. Lyman and I. Levental, "Plasma membranes are asymmetric in lipid unsaturation, packing and protein shape," *Nature Chemical Biology* **16**, 644 (2020).
- [26] <https://github.com/ZijunH/Roche> Biochemical Pathway Generator
- [27] M. Pigliucci and M. Boudry, "Why machine-information metaphors are bad for science and science education," *Science and Education*, **20**, 453 (2011).
- [28] T. Wilhelm, "The smallest chemical reaction system with bistability," *BMC Systems Biology* **3**, 90 (2009).
- [29] J. Keener and H. Sneyd, *Mathematical Physiology*, 2nd ed., (Springer, 2008).
- [30] D. Snoke, "Systems biology as a research program for intelligent design," *Bio-Complexity* **3**, 1 (2014).
- [31] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, (CRC Press, 2006).
- [32] *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*, Z. Szallasi, J. Stelling, and V. Perival, eds. (MIT Press, 2006).