Chapter 2

Physical constraints on computation

In this chapter we briefly review some of the important fundamental constraints that physical law places on computational capabilities. These constraints will serve as the basis for the arguments in chapter 6, which will establish that reversible models are necessary for permitting the maximum possible computational power in the limiting technology.

Most of existing computer science theory today deals not with physics, but with abstract realms of pure mathematics, exploring a plethora of different models of computation having wildly varying capabilities. Sometimes these theoretical models have capabilities substantially different from those of physics as we know it.

But real-world computers are physical devices, and their ultimate potential capabilities are defined not by some arbitrarily-chosen model, but rather by the hard facts of physical law. Unfortunately, physics is not yet completely understood (witness the lack of an accepted unification of quantum mechanics with general relativity), and even those parts that *are* well understood are not usually described in terms that facilitate the use of physics itself as a model of computation.

However, physics does constrain information processing in a number of important ways that can already be identified with fairly high confidence.

2.1 Propagation speed limits

The most obvious physical limit important to information processing is the lightspeed bound for the speed at which information may propagate through space.

Physical dynamics, as currently understood, proceeds purely through *local* interactions; there is no "action at a distance." Even gravity, thought by Newton to be an instantaneous force, is now understood, in the context of general relativity, to propagate through space at only the speed of light, $c \approx 3 \times 10^8$ m/s.

Even the quantum-mechanical systems that are sometimes interpreted as demonstrating "spooky action at a distance" (such as separated EPR pairs), can be explained instead in terms of local interactions. As we will review in chapter 4, quantum dynamics is based on an "amplitude function" which is a function of the global state of a system (the whole universe if you like). This leads to a statistical behavior that may at first *appear* to require nonlocal interactions, but the wavefunction actually evolves over time through a transformation (the Hamiltonian) that can be expressed as a composition of interaction terms that are entirely spatially local.

In general, due to the locality of underlying physical law, *all* influences are restricted to traveling, at most, at the speed of light. Thus, the physical transmission of information in a computer is limited to this speed as well.

This bound is "tight" in the sense that it is, of course, already achieved in practice in our ubiquitous telecommunication systems, and in optical interconnection networks in some computers. Signals in typical electrical transmission lines travel a bit slower, about half the speed of light. But propagation times are still linear in the distance traveled.

One important exception is that signals in low-inductance, resistive wires (such as the wires on integrated circuit chips) do not actually travel at constant speed, but rather, for long wires, require propagation time that is proportional to the *square* of the length ℓ of the wire, in accordance with the diffusion equation. This unfavorable scaling presents problems in integrated circuit design today. However, even with current technology, this ℓ^2 scaling is not inevitable, but can be avoided through simple schemes such as periodic re-buffering of the signal.

2.2 Information density limits

Another important constraint for computation results from physical limits on the amount of information that can be stored within a given volume of space (such as memory in a computer). We can say with confidence that some such bounds do exist, but unfortunately their exact value is hard to determine. However, these bounds will be very important in our later arguments about the advantages of reversibility, so we will now take some time to look at the various possible answers in some detail.

Fundamental quantum mechanics appears to dictate a particular finite upper bound on the total amount of information (including entropy) that can be contained in any system, as a function of the system's physical volume and the amount of energy it contains. By the amount of information in a system, we mean simply the logarithm of the number of states that the system could occupy, given some definition of what constitutes "the system." (See §2.5.2.) According to Margolus (1996, [118]),

> [The question of the number of states] is really a very old question: the correct counting of physical states is the problem that led to the introduction of Planck's constant into physics [137], and is the basis of all of quantum statistical mechanics. The question can be answered by a detailed quantum mechanical counting of distinct (mutually orthogonal) states. It can also be well approximated in the macroscopic limit [87, 184] by simply calculating the volume of phase space accessible to the system, in units where Planck's constant is 1.

Let us look at some particular information density bounds in more detail.

2.2.1 Entropy bounds from black hole physics.

Some particular upper bounds on information content as a function of system size and energy are given by Bekenstein (1984, [15]) and by Joos and Qadir (1992, [88]). Bekenstein's bounds, which originally came out of his studies of the entropy of black holes (*e.g.*, [14]), are fairly loose, in the sense that his bounds may conceivably be much higher than the maximum information content for systems *other* than black holes. One bound Bekenstein gives ([15], eq. 1) is:

$$S < 2\pi ER/\hbar c, \tag{2.1}$$

where S is the capacity for entropy or information (in natural log units or *nats*), E is the total energy (including rest mass-energy) in a system, and R is the radius of the system.

The maximum mass-energy for a system of given radius is of course achieved only in the case of black holes, since anything with a black hole's mass within a black hole's radius has such a high surface gravity that it *is* a black hole. The radius of a black hole is proportional to its mass M according to $R = 2GM/c^2$ (Wald 1984 [183], p. 124, eq. 6.1.45), and so the energy of any system of that radius is bounded by the black hole rest energy,

$$E \leq \frac{c^4}{2G} R \tag{2.2}$$

where G is Newton's gravitational constant, $G = 6.67259 \times 10^{-11}\ \mathrm{N\,m^2/kg^2}$.

Combining (2.1) and (2.2), we have

$$S < \frac{\pi c^3}{\hbar G} R^2. \tag{2.3}$$

In other words the entropy of a system is ultimately bounded in proportion to its minimal surface area, rather than to its volume! This is somewhat counter-intuitive. Perhaps one way to understand this result is the following: Imagine growing a black hole up from a tiny size by throwing objects into it. Due to a gravitational time dialation that stretches to infinity as objects approach the horizon, items we throw at the hole never actually enter it from our point of view; the objects just keep getting closer and closer to the event horizon "surface" of the hole (and the light from them gets more and more red-shifted). From that point of view, all the information about everything we throw into the hole is held at the hole's surface. So it is perhaps understandable that the horizon area should scale up in proportion to the amount of information held there.

If this is indeed the case, the information density at the event horizon given by Bekenstein's bound is truly enormous: 1/4 nat of entropy for each square of area that is 1 Planck length, or $\ell_{\mathrm{P}} = \sqrt{G\hbar/c^3} = 1.62 \times 10^{-35}$ m, on a side. That is, an astounding 2.21×10^{70} bits per square meter, or 2.21×10^{50} bits per square Ångstrom (roughly atom-size) area. (It's probably safe to say that DRAM densities won't reach *that* level for a while!)

In any case, black holes are certainly not a very good place to store information that we might want to retrieve later, although they might conceivably be a good place to dump unwanted entropy. Macroscopic black holes have intrinsic temperatures near absolute zero, and in contrast to most systems, they get *cooler* as you dump more energy and entropy into them! (*Cf.* eq. 26 in Smith's paper [152], and his references to Hawking, his source.) So a black hole would be a sort of natural heat sink, cooler even than the cosmic microwave background which is at $\sim$3 K. But for the foreseeable future, black holes will remain rather hard to come by, so it behooves us to also consider where we stand without them.

2.2.2 Entropy bounds for a photon gas.

Much tighter bounds can be given for the entropy of normal (non black-hole) systems, given additional assumptions about their composition. This is done in Bekenstein's paper [15], as well as in papers by Joos and Qadir [88] and Smith (1995 [152]) and the related literature. Smith argues that for high-temperature systems (above 1000 K or so, roughly the melting point of ordinary solids), the maximum entropy density for a *given mass density* is approximately achieved (within a small constant factor) by a thermal photon gas, in which the entropy density (entropy per unit volume) is ([152], eq. 22)

$$\frac{S}{\mathcal{V}} = \frac{16\sqrt{\pi}}{3 \cdot 60^{1/4}} \left(\frac{c}{\hbar} \cdot \frac{M}{\mathcal{V}} \right)^{3/4} \tag{2.4}$$

where $M/\mathcal{V}$ is the energy density of the photon gas, in mass units.

This equation would appear to allow arbitrarily high entropy densities to be achieved by raising the temperature and mass-energy density, except that actually of course the energy density is itself limited as a function of a system's size, since beyond a certain point the system would form a black hole.

A full general-relativistic analysis of the situation would of course be very complex, but as a simplifying first approximation, let's derive the maximum entropy density of a sphere of photons without taking GR into account, except in the sense of allowing it to set a maximum energy for a system of given radius.

Using eq. (2.2) we find that a sphere of photons of (nonrelativistic) volume $\mathcal{V} = \frac{4}{3}\pi R^3$ must have energy density

$$\frac{E}{\mathcal{V}} \leq \frac{3\,c^4}{8\pi\,G\,R^2} \tag{2.5}$$

i.e., mass density

$$\frac{M}{\mathcal{V}} \leq \frac{3\,c^2}{8\pi\,G\,R^2} \tag{2.6}$$

to avoid gravitational collapse. Substituting this into eq. (2.4), we find that non-black-hole entropy is bounded by

$$S \leq \frac{64}{9 \cdot 60^{1/4}} \left(\frac{3\pi}{8}\right)^{3/4} \frac{c^{9/4}}{(\hbar G)^{3/4}} R^{3/2} \tag{2.7}$$

$$\approx 2.889 \left(\frac{R}{\ell_{\mathrm{P}}}\right)^{3/2} \tag{2.8}$$

where ℓ_{P} is again the Planck length. This bound, interestingly, scales with increasing radius even less rapidly than in the black hole case, where we had $S = \pi(R/\ell_{\mathrm{P}})^2$. Incidentally, the ratio $S_{\mathrm{BH}}/S_{\mathrm{PG}}$ between black-hole entropy and maximum photon-gas entropy in this (admittedly simplistic) analysis is

$$\frac{S_{\mathrm{BH}}}{S_{\mathrm{PG}}} \approx 1.087\sqrt{\frac{R}{\ell_{\mathrm{P}}}}. \tag{2.9}$$

This ratio is required to be ≥ 1 by Bekenstein's argument that black holes always maximize entropy, which implies $R \gtrsim 0.846\,\ell_{\mathrm{P}}$, perhaps suggesting that $R \approx 0.846\,\ell_{\mathrm{P}}$, or thereabouts, is a minimum physical length in some sense. The entropy of either a black hole or maximum-energy photon gas sphere having that radius is ~ 3.24 bits, suggesting perhaps there is an absolute upper limit to entropy density for objects of *any* size, on the order of ~ 1.28 bits/ℓ_{P}^3, or $\sim 3.03 \times 10^{98}$ bits/cm^3. Of course this density would not be achievable for any object greater than about a Planck length in size.

But even for larger objects, our new bound (2.8), though lower than before, is still extremely high; for example, an object as large as the Earth could still have an average entropy density of 117 Gigabytes of information per cubic Ångstrom without exceeding this limit.

Given this, it is probably necessary to back a little further away from fundamental theoretical arguments, if we want to achieve any sort of meaningful bound.

2.2.3 More reasonable mass densities.

One observation is that if we rule out burying our computers inside star-sized or at least planet-sized masses of gravitating material, then there is probably no way to apply enough pressure to get their average mass density to be much greater than in ordinary solids. At a mass density of 10 g/cm^3 (about that of lead), eq. 2.4 gives only 36 kilobytes per cubic Ångstrom.

However, achieving even this much more reasonable entropy density using photons requires extremely high temperatures. The temperature of a photon gas of energy density ρ_{E} is (solving Smith's [152] eq. 13 for T):

$$T = \sqrt[4]{\frac{\rho_{\mathrm{E}} c}{4\sigma_{\mathrm{SB}}}} \tag{2.10}$$

where σ_{SB} is the Stefan-Boltzmann constant,

$$\sigma_{\mathrm{SB}} = \frac{\pi^2}{60} \frac{k_{\mathrm{B}}^4}{c^2 \hbar^3} \tag{2.11}$$

Achieving a mass density of 10 g/cm^3 (That's one heavy field of light!) thus requires a temperature of roughly 10^9 Kelvins. It is difficult to see how such high temperatures can possibly be maintained at ordinary pressures without completely destroying any structure the computer might have. Thus we need to move to still more conservative estimates.

2.2.4 More reasonable temperatures.

For example, at a more feasible temperature such as the melting point of copper, 1356 K, the energy density of light is only 2.56×10^{-9} J/cm^3, so it is essentially weightless, at 2.85×10^{-23} g/cm^3. Moreover, the *entropy* density is then only 0.74 bits per cubic micron. (This makes sense since ordinary visible light, emitted by glowing but still-solid blackbodies, has wavelengths on the order of a micron.) This clearly is much less than the entropy density of the hot copper atoms themselves, which exist at a density of roughly 0.08 atoms per cubic Angstrom. So at temperatures where any useful solid structure can exist, the energy density of light is very low, and it is also far from maximizing the possible entropy density.

What is the entropy density in an ordinary solid material like lead? It may conceivably be on the order of kilobytes per atom, as we calculated above for light of the same density, if information about the nuclear structure is included in the count—after all, Smith's argument tells us to count the total mass-energy of the system in computing his bound.

However, without a reliable way to probe the structure of nuclei, most of this information, even if it is there, will be inaccessible as a place to store information for later retrieval. The nucleus may nevertheless be capable of absorbing some amount of thermal information (heat entropy), but I have not researched whether a figure as high as 36 kilobytes of entropy at normal temperatures is consistent with what is known about nuclear structure and the heat characteristics of atomic materials. If variability in the nuclear structure does not make a large contribution to total atomic entropy, then the actual maximum entropy per atom in normal solids is probably much lower than the 36 kilobyte figure.

However, at room temperatures, entropy density is probably not much lower than on the order of 1 bit per atom (or per cubic Ångstrom), since at those temperatures atoms have enough energy to jiggle around a little, and so will have on average 1 nat (k_{B}) of entropy ($k_{\mathrm{B}}T$ energy) per vibrational degree of freedom. For three-dimensional vibrations, there are six degrees of freedom, three of position and three of momentum, so this gives 6 nats$/\ln 2 \approx 8.66$ bits per atom. There should also be entropy contributions from variability in the nuclear spin orientation, and from electrons that are free to roam in molecular orbitals or in conduction bands. But most atoms are somewhat larger than 1 $\mathrm{\AA}^3$ in volume, so 1–10 bit$/\mathrm{\AA}^3$ is still probably the right overall order of magnitude for entropy density in normal materials.

A more detailed (but still fairly crude) analysis based on actual thermochemical data from the CRC handbook [107] suggests that experimentally, at atmospheric pressure, the entropy density for copper is indeed found to be in the rather narrow range 0.5-1.5 bits$/\mathrm{\AA}^3$ for a wide range of temperatures from room temperature up to its boiling point, and moreover that the entropy densities in a variety of other pure elemental materials are also close to this level. For mixtures, we would expect the entropy density to be potentially greater, due to the additional degree of freedom implicit in choosing what species of atom resides at any given location.

Table 2.1 summarizes the above results by giving the average entropy density of a sphere of radius 1 meter that contains the maximum entropy according to the various

Material	Upper bound on entropy density	Caveats
Black hole	4.14×10^{39} b/Å^3	Need mass ≈ Saturn; can't get info. out
Non-black hole	1.53×10^{22} b/Å^3	Requires nearly as much mass
Normal density	$\sim 3\times10^{5}$ b/Å^3	May require billion-degree temperatures
Atomic matter	$\sim$ 1–10 b/Å^3	Hand-waving estimate.

Table 2.1: Theoretical limits on entropy or information density for a 1-meter-radius sphere, in various scenarios. The radius is important because in the high-gravity regime, the maximum average entropy density decreases with increasing size. It is difficult to know which of these limits, if any, might someday be approachable in real computational systems.

bounds. (Keep in mind that entropy actually scales less rapidly than volume for the systems near black-hole mass.)

This concludes our discussion of information density limits. Although we were unable to determine precisely the maximum density that was possible, we saw that entropy density does appear to ultimately be limited by some function of energy density, such as in eq. (2.4). Furthermore, much of a system's rest mass-energy may not count for purposes of this calculation, if it is energy that is tied up in an inaccessible nucleus, for example. At this stage I believe it would be premature to predict that a density greater than say $\sim$ 10 bits per cubic Ångstrom could ever actually be achieved for stable, retrievable storage of information. I would need more information before I could make a similar statement regarding thermal entropy densities.

2.3 Information flux rate limits

Another physical quantity of importance in computation is the maximum flux (rate of flow per unit area) of information or entropy through any surface in the computer. We should point out that one class of bounds on this quantity immediately follows from the bounds of sections 2.1 & 2.2, as follows.

Suppose a material having entropy density ρ_S passes through of surface at velocity v. Then the entropy in that material is crossing the surface with exactly the flux $F_S = \rho_S v$. Section 2.2 gave us bounds on the maximum value of ρ_S, and the maximum v is of course c, so this leads immediately to corresponding bounds on F_S.

One caveat is that in normal materials traveling at near the speed of light, the relativistic length contraction of the material should increase its effective entropy density, according to

$$\rho'_{\mathrm{S}} = \gamma \rho_{\mathrm{S}} \tag{2.12}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{2.13}$$

is the normal relativistic correction factor (*cf.* [52]). We assume that a given chunk of material maintains the same entropy at high speed, but is compacted into a shorter space. This would seem to allow arbitrarily high fluxes to be attained.

However, in addition to this compression into a smaller volume, the chunk will also have its mass-energy increased (in the reference frame of the stationary surface through which the material is passing) by another factor of γ, so that the mass-energy density of the moving material will actually scale as

$$\rho'_{\mathrm{M}} = \gamma^2 \rho_{\mathrm{M}}. \tag{2.14}$$

and so, solving (2.14) for γ and substituting into (2.12), the entropy density actually scales as

$$\frac{\rho'_{\mathrm{S}}}{\rho_{\mathrm{S}}} = \sqrt{\frac{\rho'_{\mathrm{M}}}{\rho_{\mathrm{M}}}}; \tag{2.15}$$

i.e., the increase in entropy density only scales as the *square root* of the energy density. So asymptotically, we could thereby do no better than with light, which already travels at lightspeed and where the entropy density scales with the energy density to the 3/4 power, according to eq. (2.4). So the maximum entropy flux we derive from our entropy density bounds is *not* exceeded in materials traveling at relativistic speeds, if the energy invested in accelerating the material to that speed is taken into account.

Smith 1995 [152], p. 6, eq. 7 gives an explicit formula for the maximum entropy flux F_{S} using light, given an energy flux F_{E}:

$$F_{\mathrm{S}} \leq \frac{4}{3} \sigma_{\mathrm{SB}}^{1/4} F_{\mathrm{E}}^{3/4} \tag{2.16}$$

This is the formula for the entropy flux emitted by a blackbody that is at the appropriate temperature to emit energy flux F_{E}. As Smith points out, there is a simple proof that this is the maximum entropy flux that can be transmitted with photons given that energy flux. Imagine using photons to continuously transmit energy and entropy through a small aperture into an insulated box (a perfect blackbody). The interior of the box will heat up, and, at equilibrium, will radiate energy out of the aperture exactly as fast as it is coming in (since energy is conserved), and will also radiate entropy out *at least* as fast as it comes in (since global entropy cannot decrease). Therefore the entropy flux of the thermal blackbody radiation coming out of the box upper-bounds the achievable entropy flux of the light coming in, which may be of any form (coherent, *etc.*).

At this point we could go on to calculate upper bounds on information flux at *any* energy based on the black hole limits to entropy density that we discussed in §2.2. For example, for a postulated minimum-size (Planck-length scale) black hole moving at near the speed of light, we estimate entropy flux would be around 10^{109} bit/s-cm^2. However, this sort of bound is rather far from anything meaningful, since it does not represent a sustainable rate, or a rate achievable over an area much larger than a Planck length—black holes placed near each other would rapidly conglomerate into a larger black hole with lower entropy density. Even if we were so bold as to allow for the use of such exotic objects as black holes as computer components, properly accounting for gravitational effects in such systems would make our scaling analysis of chapter 6 much more complex. So instead, for the rest of the thesis, we will ignore high-gravity situations, and instead focus only on the bounds obtained for normal matter.

2.4 Computation rate limits

In chapter 6 we will examine in detail how certain kinds of limits on computation rates for irreversible and imperfectly-reversible computers can be derived from the limits on information flux we saw in §2.3. However, there are other limits on processing rates that apply even to perfectly reversible computers.

In particular, there is the result of Margolus and Levitin (1996, [118]) that the fundamental laws of quantum mechanics imply that the maximum rate $\nu_\perp$ at which a system at an average energy E (above some minimum energy E_0) can transition between distinguishable (*i.e.*, orthogonal) states is

$$\nu_\perp \leq 4(E - E_0)/h. \tag{2.17}$$

This bound is derived in a totally general way, and applies even for systems traveling at relativistic velocities. Insofar as any computational operation requires that some part of a system change from one distinct state to another, Margolus and Levitin's bound is an absolute upper limit on the rate at which operations can be performed within a computer.

Further, Margolus suggests [personal communication] that for systems in which not all the system's energy is accessible for computational purposes (for example, if some of it is in the form of heat, or tied up in rest mass), it is the *free energy* of the system, rather than its total energy, that determines the maximum rate at which the system can transition between useful computational states according to eq. (2.17).

As a simple example, a single electron excited to a potential of 1 Volt above its ground state contains 1 eV of accessible energy and thus can never perform computational steps (or any state change) more rapidly than at a rate of 4 eV/h = 9.67×10^{14} Hz, or about once per femtosecond.

2.5 Reversibility of physics

Another physical constraint of great importance for computation is that all physical dynamics is reversible (invertible), that is, it is deterministic looking backwards in time. (See figure 2.1.)

Quantum mechanics is sometimes described in nondeterministic terms, but it is actually perfectly deterministic (and reversible) at the level of the evolution of the quantum wave function. The apparent nondeterminism of quantum events can be interpreted as merely a subjective, emergent phenomenon that is predicted perfectly well by the underlying deterministic theory [57].

One possible exception to reversibility may be black holes, which, in some theoretical arguments, are found to destroy information (see Preskill 1992 [138] for a review of the situation). However, there is currently no accepted, complete theory of black hole physics from which we could draw indisputable theoretical conclusions, and there is no experimental evidence that supports information loss. The truth of the issue is still being actively debated (*e.g.*, [55, 120]). Moreover, it appears that some recent developments in string theory would allow reversibility to be maintained, if the theory is correct (Myers 1997, [130]).

In any case, it seems to be the general consensus among physicists that reversibility is certainly maintained in at least all areas of mechanics that do not involve extreme situations such as black holes. So regardless of the black hole situation, physics remains reversible for all practical purposes.

2.5.1 Physical reversibility and information erasure

Another way of characterizing physical reversibility is that two states (of a classical system, or of a quantum wavefunction) that are initially distinct can never evolve to become the *same* state at some later time. (In the language of functions, the system's transition function over any time period is one-to-one/bijective/invertible.) Consequently, the number of possible states of a system is irreducible over time; we say that the system's state space is incompressible. The invertibility is a simple consequence

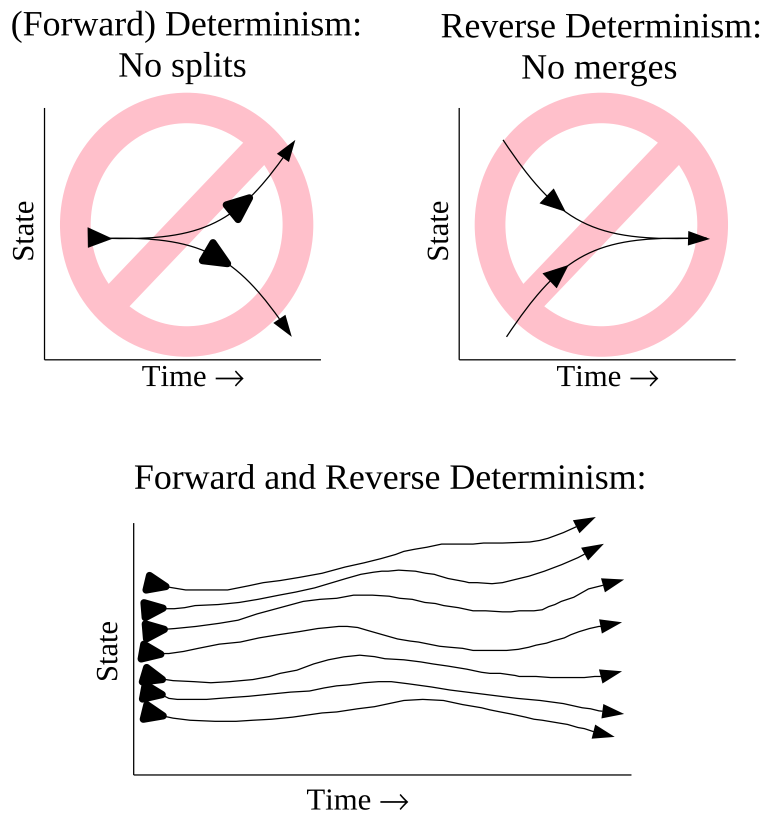

Figure 2.1: Forward and reverse determinism in physics. Normal forward determinism means that a single state cannot evolve to become one of two different states at any single later time, and similarly, reverse determinism or just *reversibility* means that two initially-distinct states cannot evolve to become the *same* state at some later time. Physics is both forward and reverse deterministic, and so the possible trajectories of a system through configuration space-time never intersect.

of the fact that in any standard description of physics, such as Hamiltonian dynamics, an isolated physical system evolves according to a time-differential equation. The differentials apply equally well in either time direction.

[Note: Our "incompressibility of state space" concept is closely related to, but not the same thing as, another property that is shared by all systems with Hamiltonian dynamics, namely that their "phase space *volume*" is incompressible. However, we will not delve into those rather subtle distinctions here, as they are not important for our immediate purposes.]

In any event, the incompressibility of state space has an important consequence for information erasure within a computer, first described explicitly by Landauer [97]. Whenever we attempt to irreversibly erase a piece of information from a computer, that information is not truly destroyed, but instead is simply *transferred* to another part of the system, typically to the uncontrolled thermal state of the computer and its environment. We explain in more detail with reference to figure 2.2.

The figure illustrates a 1-bit piece of computational state within a computer. We wish to perform an "erasure" operation, which we may characterize as an operation that transforms that bit to a zero regardless of whether it was originally a 0 or a 1.

In addition to the bit in question, the computer also contains some amount of other information in the form of other bits in memory, together with the entropy of its thermal state. Let $\mathcal{N}$ denote the number of possible states of the system, apart from the bit in question.

We want our "erase" operation to operate correctly, independently of which of the $2\mathcal{N}$ possible states the combined system is in. Due to physical reversibility, each of these $2\mathcal{N}$ states must be mapped to a distinct state after the erase operation—but all of those states have value 0 in the erased bit. Thus there must be $2\mathcal{N}$ possible states of the rest of the system, after the operation. The amount of information in the rest of the system has therefore increased by $\lg 2\mathcal{N} - \lg \mathcal{N} = 1$ bit.

So the presumably erased information has not really been destroyed, but is still present somewhere, either in some other part of the computational state or in the thermal state. The original value of the bit could *in principle* be retrieved by, for example, running the laws of physics backwards.

However, if the information has been lost in a sea of thermal chaos, then *in practice* there is *no* way to reconstruct the original value of the bit.

2.5.2 Reversibility, entropy, and the second law

We now see how physical reversibility can be understood to imply the second law of thermodynamics.

The second law of thermodynamics states that the total entropy of any closed system cannot decrease. What is entropy? Quantitatively, it is the logarithm of the

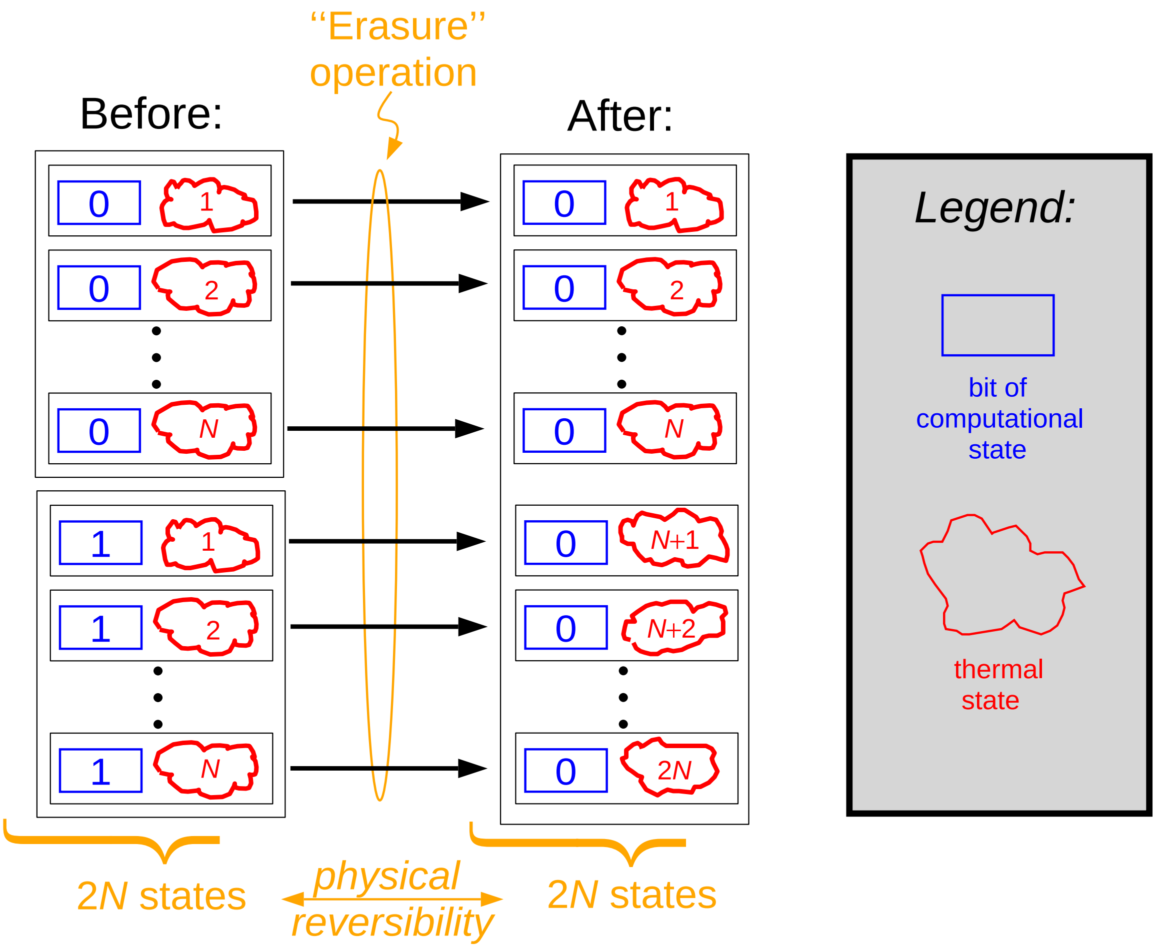

Figure 2.2: Information "erasure" under reversible physics. In order to erase an unknown bit and thereby reduce the number of possible digital states of the computer by a factor of 2, one has to make up for this by increasing the number of possible thermal states of the rest of the system by a factor of 2.

number $\mathcal{N}$ of possible states of a system. The base of the logarithm determines the unit of entropy: if the base is $e \approx 2.718\ldots$, the base of natural logarithms, then we might call the corresponding unit of entropy 1 *nat*, equal to Boltzmann's constant k_B. If the base is 2, then the unit of entropy is called 1 *bit*. Thus, 1 bit $= (\ln 2)$ nat $\approx$ 0.693 nat.

If entropy is the log of the number of possible states, what, then, do we mean by a "possible state?" This depends entirely on the context, specifically on how we define what constitutes a legal example of "the system" in question.

However, even with this broad definition of entropy, we can already make some meaningful statements about entropy in connection with reversibility. First, it is clear that if the entropy of a system were to decrease over time, then the system would not be reversible, because we would have an example of multiple possible initial states evolving to become a smaller number of resulting final states, violating the incompressibility of state space that is implied by reversibility. Therefore, the reversibility of a system immediately implies that its entropy can never decrease over time.

There is a similar connection between entropy increase and determinism. In a deterministic system, state space is "inexpandable" since a given state can not evolve to more than one possible new state in a given amount of time. Thus the number of possible states cannot increase, in this strict sense, and so deterministic systems undergo no "true" increases in entropy.

However, even if a system is deterministic, we may find it convenient to label more and more states as "possible" over a system's time evolution, simply because, given an incomplete model of a system's initial state, we may lose track of the exact trajectories of the initially possible states over time, and so many additional states may become possible over time from the point of view of the model. In such circumstances, it is convenient to say that entropy increases. An example is the situation in figure 2.2 (p. 44). Suppose we have constructed an initial condition in which only the "1" value of the bit is possible, so the entropy before the "erase" operation is just $\ln \mathcal{N}$. But since we fail to model what becomes of the information *that* the bit is 1 after the "erase" operation is performed, the entropy of the system under the model increases to $\ln 2\mathcal{N}$. This increase will happen whenever a non-entropy bit turns into thermal form, because the evolution of the micro-state of a thermal system is, by definition, un-trackable by us.

We thus can state the following principle: Total entropy increases (permanently) by at least 1 bit's worth any time a bit that is originally non-entropic moves to reside in a thermal system. Furthermore, this happens whenever a digital bit is erased, unless (a) the bit was *already* entropy, in which case moving it to thermal form does not necessarily increase total entropy, or (b) the bit is canceled out instead, by un-computing it from other bits of state with which it is correlated.

2.5.3 Entropy and energy

As we just saw, the second law of thermodynamics states that the entropy of a closed system cannot decrease over time. We saw that it also cannot increase, except in the sense that our incomplete model of the system may lose track of what happens to a state over time, so that more states become "possible" from the point of view of the model. However, entropy *can* be moved from one subsystem to another.

Correspondingly, the *first* law of thermodynamics states that the *energy* of a closed system can neither increase nor decrease, but can only be moved from one subsystem to another.

How are these conservation laws for energy and entropy related? We find empirically that in order to increase or decrease the entropy of any subsystem (not counting increases due to deficiencies in our model), we generally must also increase or decrease its energy (given a closed, constant-volume system). For sufficiently small changes in entropy, we find that the change in energy required is proportional to the change in entropy. The constant of proportionality is called the *temperature* T of the subsystem. Formally,

$$T = \partial E / \partial S. \tag{2.18}$$

This is a perfectly valid definition of temperature, in terms of the relation between a system's energy and its number of states.

Under this definition, 1 Kelvin of absolute temperature is definable as a requirement of $\approx 1.38 \times 10^{-23}$ J of energy per 1-nat increase in entropy. A nat of entropy can therefore also be expressed in units of energy per unit temperature, such as 1.38×10^{-23} J/K. In such form, 1 nat of entropy is often referred to as *Boltzmann's constant* k_{B}.

From all this, it follows immediately that the amount of energy E that must be added to system in order to double its number of possible states is just

$$E = k_{\mathrm{B}} T \ln 2 \tag{2.19}$$

since $k_{\mathrm{B}} \ln 2$ is just a 1-bit increase in entropy, and multiplying by the system's temperature just converts this entropy increase to the required change in energy, by the definition of temperature.

2.5.4 Logical irreversibility and energy dissipation

We saw in section 2.2 that in any system with particular size and energy there is a consequent upper bound on the entropy that system can contain. If a given system is found to contain less entropy than the maximum given the amount of energy in the

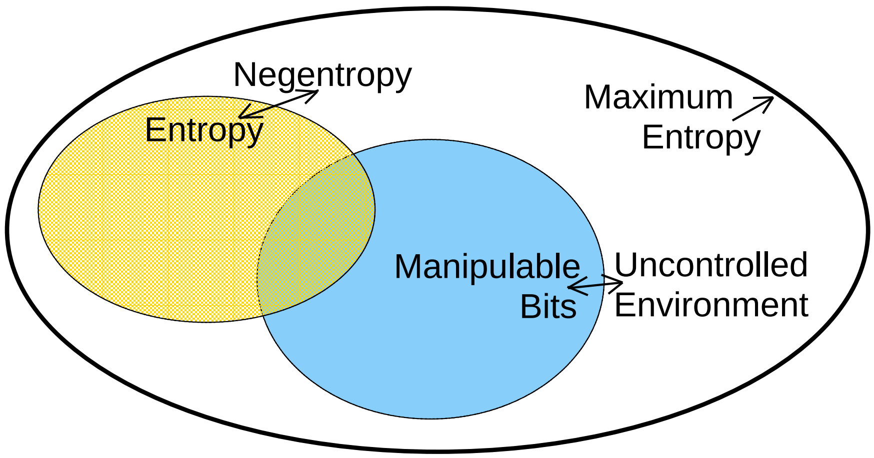

Figure 2.3: Venn diagram of entropy and information. Any system of finite size and energy has a finite maximum entropy; however, if the system expands without bound, the maximum entropy may also. The maximum entropy may be considered the total amount of information of all kinds in the system. However, much of it may be redundant and cross-correlated. Bits of information that are uncorrelated, or whose correlations have become lost beyond all hope of recovery, are entropy, which can only increase in a reversible universe. Some portion of the system, and the information in it, is within our ability to manipulate and control, such as bits within a computer. These bits, too, may either be entropy or not, depending on our ability to know their correlations.

---

system, then that must mean our model of the system is imposing further structure on the system, ruling out some of the states that would otherwise be possible.

For such a system with non-maximal entropy, only a portion of the energy of the system is actually needed for permitting the entropy that is actually present. This portion of the total energy will be referred to as the amount of *dissipated energy* in the system. The rest of the system's energy will be referred to as its *free energy*. The difference between the entropy of the system and its maximum entropy will be termed the *negentropy* or *information capacity* of the system. Some of this information capacity may become allocated for storing computational information. (See fig. 2.3.)

As we discussed in §2.5.2, even in the context of a perfectly deterministic underlying physics, the entropy of a system can be seen to increase, through a failure to completely model the determinism inherent in the system's physical evolution. When this happens, the amount of the system's energy that is needed to support this entropy will increase by some amount, and the free energy will decrease by the same amount. We say this amount of energy has been *dissipated*.

We are now in a position to accurately state and explain the central statement on which the field of reversible computing is based:

Landauer's principle. *The irreversible loss of 1 bit of computational information requires the dissipation of* $k_B T \ln 2$ *energy*, where T is the temperature of the subsystem in which the lost bit finally ends up. By the "irreversible loss," we mean that some bit of *computational information* (not a bit that is already entropy!) becomes transformed in such a way that our computational models can not track it, for example by becoming mixed up with parts of the system whose state is already *thermal*, or unknown. Thus by definition the bit has become entropy, and the entropy of the system as a whole is increased by 1 bit. This increase is eventually reflected in some subsystem at temperature T, and by definition of temperature, the energy of this subsystem must be increased by $k_B T \ln 2$. The energy invested in the entropy increase is heat. If T is the lowest available temperature, then this energy must come out of the free energy, because all the dissipated energy in the system is already fully occupied with containing the pre-existing entropy. Thus the free energy is decreased by $k_B T \ln 2$.

The above principle was first explicitly conjectured by Landauer [97].

Note that since T is the temperature of the system where the entropy finally ends up, not the temperature of the device that held the entropy originally, cooling a computer cannot in the long run decrease the total energy dissipation required to erase bits, if the dissipation in the cooling system is taken into account. The entropy that is generated can not build up indefinitely in the cooling system, or else it would not stay cool. Instead, it ultimately ends up in some natural thermal reservoir in the environment. The coolest thermal reservoir of effectively unlimited capacity that might be available in the foreseeable future is the interstellar microwave background, at a temperature of $\sim$2.73 K. Thus, no process that generates entropy can, in the long run, sustain an energy dissipation cost less than $k_B(2.73\,\mathrm{K}) \ln 2 \approx 2.6 \times 10^{-23}$ J per bit generated, and this can only be attained if the entropy can be transmitted directly into space. For terrestrial systems that use the atmosphere as their thermal reservoir, the relevant temperature is in the neighborhood of room temperature or 300 K, for a minimum energy dissipation of $\sim 3 \times 10^{-21}$ J/b.

2.6 Quantum computation

One area in which physics may actually constrain computation *less* than might be expected is in the possibility of quantum computation (*cf.* [60, 47, 23, 22, 21, 149, 146]), that is, computation using large, complex, coherent superpositions of states. If it can be implemented successfully, quantum computation seems likely to be strictly asymptotically faster than classical computation on certain problems, by as much as an exponential factor. But it is not yet known if quantum computation would be beneficial for purposes other than obsoleting the RSA cryptosystem, or simulating physical quantum systems. Still, if the recent progress on implementing quantum computers [56, 168, 35, 39, 76, 40, 158, 49, 150, 10, 34] eventually culminates in success, then we would certainly like to consider quantum computation as a physically possible means of computation. But even a quantum computer would still need to obey the fundamental constraints discussed above affecting the maximum density and propagation speed of information.

We will review quantum computing in more detail in chapter 4.

2.7 Physical constraints—conclusion

This concludes our discussion of fundamental physical limits on computation. Table 2.2 summarizes the limits we discussed, and the presumed effect on the form of a physically-realistic model of computation, which we will discuss further in ch. 5.

In chapter 6 we will see how these limits affect the scaling of computation speeds in reversible and irreversible computers. But first, in the next chapter, we review the non-physical theoretical underpinnings of reversible computing, and show that in an imagined non-physical computational framework, reversibility leads to unfavorable scaling. The contrast between that result and the results of chapter 6 underscores that traditional non-physical theoretical frameworks for computation are inadequate for realistically modeling the advantages of reversibility, and thus, more sophisticated models of computation that take the above-described physical constraints into account are required for a correct analysis. Such models will be discussed in chapter 5.

Fundamental principle	**Constrained quantity**	**Symbol**	**Quantitative constraint**	**Impact on our model**
Quantum mechanics	Entropy density	ρ_{S}	$\lesssim$ 1–10 b/Å^3?	Finite state/ processor
	Entropy flux	F_{S}	$\leq \rho_{\mathrm{S}} v$	Finite info. flux
	Rate of state change	$\nu_{\perp}$	$\leq 4(E - E_0)/h$	Finite oper. frequency
Locality	Info. prop. velocity	v	$\leq c \approx 3 \times 10^8$ m/s	Mesh arch. (Vitányi '88)
3-dimensionality of space	Connec-tivity		$\mathcal{O}(t^3)$	3-D mesh
Micro-reversibility, thermodynamics	Entropy change	ΔS	≥ 0 always, ≥ 1 bit/bit erasure	Logical reversibility, entropy accounting
	Energy dissipation	ΔE	≥ 0 always, $\geq k_{\mathrm{B}} T \ln 2$/eras.	
Frictional effects	Entropy coefficient	k_{S}	> 0 b/Hz?	Time-prop. reversibility

Table 2.2: Fundamental physical constraints on computation, and their effects on the form of a physically-realistic model of computation. The value of the bound on ρ_{S} is very uncertain, but the assertion that some such bound exists is not. For any particular computing technology through the foreseeable future, there will generally be much stricter limits than the above on most of these quantities.