# Maxwell's Demon

Kirk T. McDonald

Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544 (October 3, 2004; updated September 20, 2016)

## 1 Problem

This discussion is based on Prob. 2 of [1].

One of the earliest conceptual "supercomputers" was Maxwell's Demon [3]-[7], who uses intelligence in sorting molecules to appear to evade the Second Law of Thermodynamics.

#### 328 Molecular Theory.

When the gas is so far condensed that it assumes the liquid or solid form, then, as the molecules have no free path, they have no regular vibrations, and no bright lines are commonly observed in incandescent liquids or solids. Mr. Huggins, however, has observed bright lines in the spectrum of incandescent erbia and lime, which appear to be due to the solid matter, and not to its vapour.

#### LIMITATION OF THE SECOND LAW OF THERMODYNAMICS.

Before I conclude, I wish to direct attention to an aspect of the molecular theory which deserves consideration.

One of the best established facts in thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or of pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the tem-

#### Statistical Knowledge of Bodies. 329

perature of B and lower that of A, in contradiction to the second law of thermodynamics.

This is only one of the instances in which conclusions which we have drawn from our experience of bodies consisting of an immense number of molecules may be found not to be applicable to the more delicate observations and experiments which we may suppose made by one who can perceive and handle the individual molecules which we deal with only in large masses.

In dealing with masses of matter, while we do not perceive the individual molecules, we are compelled to adopt what I have described as the statistical method of calculation, and to abandon the strict dynamical method, in which we follow every motion by the calculus.

It would be interesting to enquire how far those ideas about the nature and methods of science which have been derived from examples of scientific investigation in which the dynamical method is followed are applicable to our actual knowledge of concrete things, which, as we have seen, is of an essentially statistical nature, because no one has yet discovered any practical method of tracing the path of a molecule, or of identifying it at different times.

I do not think, however, that the perfect identity which we observe between different portions of the same kind of matter can be explained on the statistical principle of the stability of the averages of large numbers of quantities each of which may differ from the mean. For if of the molecules of some substance such as hydrogen, some were of sensibly greater mass than others, we have the means of producing a separation between molecules of different masses, and in this way we should be able to produce two kinds of hydrogen, one of which would be somewhat denser than the other. As this cannot be done, we must admit that the equality which we assert to exist between the molecules of hydrogen applies to each individual molecule, and not merely to the average of groups of millions of molecules.

To place the demon in a computational context, consider a computer "memory" that consists of a set of boxes (bits), each of volume V and each containing a single molecule. A (re)movable partition divides the volume into "left" and "right" halves. If the molecule is in the left half of a box this represents the 0 state, while if the molecule is in the right half of a box we have the 1 state.



The boxes are all at temperature T, as maintained by an external heat bath. By averaging over the motion of each molecule, we can speak of the pressure P in each box according to the ideal gas law, P = kT/V, where k is Boltzmann's constant.

### (a) A Model for Classical Erasure of a Bit

A memory bit can be erased (forced to the 0 state) without knowledge as to the value of that bit by the following sequence of operations:

- Remove the partition, permitting a free expansion of the gas from volume v to 2V.
- Isothermally compress the volume of the box from 2V back to V by means of a piston that moves from the far right of the box to its midplane. The molecule is now in the left half of the box, no matter in which half it originally was.
- Reinsert the partition (at the right edge of the compressed volume).
- Withdraw the piston, restoring the box to its original shape, with the molecule in the left half of the box and nothing in the right half = the 0 state.

Deduce the total entropy change of the system of memory + thermal bath for the combined processes of free expansion followed by isothermal compression.

Exercise (a) illustrates Landauer's Principle [8] that in a computer which operates at temperature T there is a minimum entropy cost of  $k \ln 2$  to perform the "logically irreversible" step of erasure of a bit in memory, while in principle all other types of classical computational operations could be performed (reversibly) at zero entropy and zero energy cost.<sup>1</sup>

An important extrapolation from Landauer's Principle was made by Bennett [9, 10] who noted that if a computer has a large enough memory such that no erasing need be done during a computation, then the computation could be performed **reversibly**, and the computer restored to its initial state at the end of the computation by undoing (reversing) the program once the answer was obtained.

The notion that computation could be performed by a reversible process was initially considered to be counterintuitive – and impractical. However, this idea was of great conceptual importance because it opened the door to quantum computation, based on quantum processes which are intrinsically reversible (except for measurement).

A second important distinction between classical and quantum computation (*i.e.*, physics), besides the irreversibility of quantum measurement, is that an arbitrary (unknown) quantum state cannot be copied exactly [15]-[19].

<sup>&</sup>lt;sup>1</sup>The principle enunciated by Landauer himself [8] is that erasure has an energy cost of at least  $kT \ln 2$ . However, the present example does not confirm this claim.

#### (b) Classical Copying of a Known Bit

In Bennett's reversible computer there must be a mechanism for preserving the result of a computation, before the computer is reversibly restored to its initial state. Use the model of memory bits as boxes with a molecule in the left or right half to describe a (very simple) process whereby a bit, whose value is known, can be copied at zero energy cost and zero entropy change onto a bit whose initial state is **0**.

A question left open by the previous discussion is whether the state of a classical bit can be determined without an energy cost or entropy change.

In a computer, the way we show that we know the state of a bit is by making a copy of it. To know the state of the bit, *i.e.*, in which half of a memory box the molecule resides, we must make some kind of **measurement**. In principle, this can be done very slowly and gently, by placing the box on a balance, or using the mechanical device (from [11]) sketched below, such that the energy cost is arbitrarily low, in exchange for the measurement process being tedious, and the apparatus somewhat bulky. Thus, we accept the assertion of Bennett and Landauer that measurement and copying of a classical bit are, in principle, energy-cost-free operations.<sup>2</sup>



We can now contemplate another procedure for resetting a classical bit to 0. First, measure its state (making a copy in the process), and then subtract the copy from the original. (We leave it as an optional exercise for you to concoct a procedure using the

 $<sup>^{2}</sup>$ However, there is a kind of hidden entropy cost in the measurement process; namely the cost of preparing in the 0 state the bits of memory where the results of the measurement can be stored.

molecule in a box to implement the subtraction.) This appears to provide a scheme for erasure at zero energy/entropy cost, in contrast to the procedure you considered in part (a). However, at the end of the new procedure, the copy of the original bit remains, using up memory space. So to complete the erasure operation, we should also reset the copy bit. This could be done at no energy/entropy cost by making yet another copy, and subtracting it from the first copy. To halt this silly cycle of resets, we must sooner or later revert to the procedure of part (a), which did not involve a measurement of the bit before resetting it. So, we must sooner or later pay the energy/entropy cost to erase a classical bit.

Recalling Maxwell's demon, we see that his task of sorting molecules into the left half of a partitioned box is equivalent to erasing a computer memory. The demon can perform his task with the aid of auxiliary equipment, which measures and stores information about the molecules. To finish his task cleanly, the demon must not only coax all the molecules into the left half of the box, but he must return his auxiliary equipment to its original state (so that he could use it to sort a new set of molecules...poor demon). At some time during his task, the demon must perform a cleanup (erasure) operation equivalent to that of part (a), in which the entropy of the molecules/computer can be unchanged, but with an increase in the entropy of the environment by at least  $k \ln 2$  per erased bit.

The erasure demon obeys the Second Law of Thermodynamics<sup>3</sup> – and performs his task millions of times each second in your palm computer.

The "moral" of this problem is Landauer's dictum:

Information is physical



## 2 Solution

### 2.1 Classical Erasure

It is assumed that the partition can be removed and inserted without expenditure of energy (without any flow of heat).

The entropy change of a gas during the irreversible free expansion is that same as that during a slow, reversible isothermal expansion between the same initial and final states.

During an isothermal expansion of volume V to 2V, the work done by the molecule is,

$$W_{\rm by\ molecule} = \int_{V}^{2V} P\ dV = kT \int_{V}^{2V} \frac{dV}{V} = kT\ln 2.$$

$$\tag{1}$$

To keep the temperature, and hence the internal energy, of the molecule constant, heat must flow into the box from the heat bath, and in amount  $Q = WkT \ln 2$ . Hence, the

<sup>&</sup>lt;sup>3</sup>For further reading, see chap. 5 of [12], and [10, 13, 14].

thermodynamic entropy change of the box during the isothermal expansion is  $\Delta S_{\text{box, iso exp}} = \Delta Q_{\text{into box}}/T = k \ln 2$ . The entropy change of the bath is equal and opposite,  $\Delta S_{\text{bath, iso exp}} = -k \ln 2$ .

During the free expansion to the same final state as of the isothermal expansion, the entropy change of the box is the same as during the isothermal expansion, but the thermal bath experiences no entropy change as it transfer no heat.

$$\Delta S_{\text{box, free exp}} = k \ln 2, \qquad \Delta S_{\text{bath, free exp}} = 0.$$
 (2)

Then, during the isothermal compression, the entropy changes of the box and bath are just the opposite of those during the isothermal expansion,

$$\Delta S_{\text{box, iso comp}} = -k \ln 2, \qquad \Delta S_{\text{bath, iso comp}} = k \ln 2. \tag{3}$$

Thus, the total entropy changes during the erasure of the bit, consisting of a free expansion followed by an isothermal compression, are,

$$\Delta S_{\text{box, erasure}} = 0, \qquad \Delta S_{\text{bath, erasure}} = k \ln 2, \tag{4}$$

and the total entropy change of the universe is,

$$\Delta S_{\text{universe, erasure}} = k \ln 2. \tag{5}$$

Note that there is no net energy change in either the free expansion or the isothermal compression, so the erasure of a bit is accomplished at zero energy cost in this model; this is, of course, a consequence of conservation of total energy in the Universe. However, the agent that performs the isothermal compression does work  $kT \ln 2$ , which is the energy cost to that agent in performing the erasure of the bit. This is the sense of Landauer's claim (footnote 1, p. 2) that there is an energy cost of at least  $kT \ln 2$  in erasing a bit.

### 2.2 Classical Copying of a Known Bit

The original bit box has its molecule in either the left half (0) or the right half (1), and we know which is the case. The copy box is initially in a particular state that we might as well take to be 0, *i.e.*, its molecule is in the left half.

The copying can be accomplished as follows:<sup>4</sup>

- 1. If the original bit is 0, do nothing to the copy bit, which already was 0.
- 2. If the original bit is 1, rotate the copy box by 180° about an axis in its left-right midplane. After this, the molecule appears to be in the right half of the copy box, and is therefore in a 1 state as desired.

No energy is expended in any of these steps. No heat flows. Hence, there is no (thermodynamic) entropy change in either the computer or in the environment.

The rotation of the copy box by 180° is equivalent to the logical NOT operation. Thus, the copying procedure suggested above could be called a **controlled-NOT** operation, in which

<sup>&</sup>lt;sup>4</sup>A scheme involving two pistons and the (re)movable partition, but no rotation, also works.

the NOT operation is performed only if a relevant control bit is in the 1 state. Since all computation involves changing 0's into 1's and *vice versa*, we get a preview of the important role of controlled-NOT operations in classical and quantum computation. Indeed, since all classical computation can be performed using logic gates based on controlled-controlled-NOT operations [20], the present example suggests that classical computation could in principle be performed at zero energy cost, and with an entropy cost associated only with erasure of the memory.

The trick of rotating a box by  $180^{\circ}$  if it is in a 1-state to bring it to the 0-state is a possible process for classical erasure of a bit. However, this process requires knowledge of the initial state of the box, whereas the method or part (a) does not require such knowledge. So, rotation of the box is not a solution to part (a) as posed.

Since the rotation of the box is a reversible process, it doesn't change entropy. Could it then be that the trick of rotating the box provides a means of erasure with no entropy cost?

The issue now whether the task of acquiring the knowledge as to the state of the both implies an increase of entropy, of at least  $k \ln 2$ .

This is a famous question, associated with the concept of negentropy – that information is associated with a kind of negative entropy, and that the creation of information implies a corresponding increase of entropy somewhere in the larger system. A sense of this was noted already in 1868 by Tait, p. 100 of [21]. Longer discussions were given by Szilard [22] and by Brillouin [23]. See also [24]-[31].

If the box has moment of inertia I and we wish to accomplish the erasure in time t, we give the box constant angular acceleration  $\ddot{\theta} = 4\pi/t^2$  for time t/2 and then the negative of this for an additional time t/2 to leave the box at rest after rotation by  $\pi$  radians. The torque required during this process is  $\tau = I\ddot{\theta} = 4\pi I/t^2$ , so the work done in rotating the box is  $W = \tau \Delta \theta = \pi \tau = 4\pi^2 I/t^2$ .

For a box that consists of a pair of C<sub>60</sub> "buckyballs," of mass  $m = 60m_{\rm C} \approx 720m_p$  and radius  $r \approx 0.5$  nm each, the moment of inertia is  $I \approx 2mr^2(1+2/3) = 10mr^2/3 \approx 6 \times 10^{-16}m_p \approx 10^{-42}$  J. If the erasure to be accomplished in time t = 0.1 ns (10 GHz), the work done in rotating the box is  $W = 4\pi^2 I/t^2 \approx 4 \times 10^{-21}$  J  $\approx 1/40 \text{ eV} = kT$  for room temperature.

That is, even the rotating pair of buckyballs as a memory element obeys Landauer's claim that the energy cost of erasure is at least  $kT \ln 2$  for "practical" parameters.

## References

- K.T. McDonald, Physics of Quantum Computation (Jan. 31, 2005), http://kirkmcd.princeton.edu/examples/ph410problems.pdf
- P.M. Harman, Ed., The Scientific Letters and Papers of James Clerk Maxwell, Vol. 2, 1862-1873 (Cambridge U. Press, 1995).
- [3] J.C. Maxwell, Letter to P.G. Tait (11 Dec. 1867), Letter 277, p. 328 of [2]. See pp. 332-332. http://kirkmcd.princeton.edu/examples/statmech/maxwell\_letter\_277.pdf
- [4] J.C. Maxwell, Letter to J.W. Strutt (6 Dec. 1870), Letter 350, p. 582 of [2]. http://kirkmcd.princeton.edu/examples/statmech/maxwell\_letter\_350.pdf

- [5] J.C. Maxwell, The Theory of Heat (Longmans, 1871), p. 308 http://kirkmcd.princeton.edu/examples/statmech/maxwell\_heat\_71.pdf
   See also pp. 213-214 of C.G. Knott, ed., Life and Scientific Work of Peter Guthrie Tait (Cambridge U.P., 1911), http://kirkmcd.princeton.edu/examples/statmech/tait\_papers\_sup.pdf
- [6] W. Thomson, *Kinetic Theory of the Dissipation of Energy*, Nature 9, 442 (1874). http://kirkmcd.princeton.edu/examples/statmech/thomson\_nature\_9\_441\_74.pdf
- [7] W. Thomson, The Sorting Demon of Maxwell, Nature 20, 126 (1879). http://kirkmcd.princeton.edu/examples/statmech/thomson\_nature\_20\_126\_79.pdf
- [8] R. Landauer, Irreversibility and Heat Generation in the Computing Process, IBM J. Res. Dev. 5, 183 (1961), http://kirkmcd.princeton.edu/examples/QM/landauer\_ibmjrd\_5\_183\_61.pdf
- C.H. Bennett, Logical Reversibility of Computation, IBM J. Res. Dev. 17, 525 (1973), http://kirkmcd.princeton.edu/examples/QM/bennett\_ibmjrd\_17\_525\_73.pdf
- [10] C.H. Bennett, Notes on the history of reversible computation, IBM J. Res. Dev. 32, 16 (1988), http://kirkmcd.princeton.edu/examples/QM/bennett\_ibmjrd\_32\_16\_88.pdf
- [11] W.H. Zurek, Algorithmic Randomness, Physical Entropy, Measurements, and the Demon of Choice, (July 1, 1998), http://arxiv.org/abs/quant-ph-9807007
- [12] R.P. Feynman, The Feynman Lectures on Computation (Addison-Wesley, 1996), http://kirkmcd.princeton.edu/examples/QM/feynman\_computation.pdf
- [13] C.H. Bennett, The Thermodynamics of Computation A Review, Int. J. Theor. Phys. 21, 905 (1982), http://kirkmcd.princeton.edu/examples/QM/bennett\_ijtp\_21\_905\_82.pdf
- [14] J. Bub, Maxwell's Demon and the Thermodynamics of Computation, Stud. Hist. Phil. Mod. Phys. 32, 569 (1961), http://kirkmcd.princeton.edu/examples/QM/bub\_shpmp\_32\_569\_01.pdf
- [15] W.K. Wootters and W.H. Zurek, A single quantum cannot be cloned, Nature 299, 802 (1982), http://kirkmcd.princeton.edu/examples/QM/wootters\_nature\_299\_802\_82.pdf
- [16] D. Dieks, Communication by EPR Devices, Phys. Lett. A 92, 271 (1982), http://kirkmcd.princeton.edu/examples/QM/dieks\_pl\_a92\_271\_82.pdf
- [17] P.W. Milonni and M.L. Hardies, Photons Cannot Always Be Replicated, Phys. Lett. A 92, 321 (1982), http://kirkmcd.princeton.edu/examples/QM/milonni\_pl\_a92\_321\_82.pdf
- [18] The essence of the no-cloning theorem had been demonstrated earlier, but was little noticed; J.L. Park, The Concept of Transition in Quantum Mechanics, Found. Phys. 1, 23 (1970), http://kirkmcd.princeton.edu/examples/QM/park\_fp\_1\_23\_70.pdf
- [19] My demonstration of the no-cloning theorem is on p. 28 of [1].
- [20] E. Fredkin and T. Toffoli, Conservative Logic, Int. J. Theor. Phys. 21, 219 (1982), http://kirkmcd.princeton.edu/examples/QM/fredkin\_ijtp\_21\_219\_82.pdf

- [21] P.G. Tait, Sketch of Thermodynamics (Douglas, Edinburgh, 1868), http://kirkmcd.princeton.edu/examples/statmech/tait\_thermo\_68.pdf http://kirkmcd.princeton.edu/examples/statmech/tait\_thermo\_68\_p100.txt
- [22] L. Szilard, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Vesen, Z. Phys. 53, 840 (1929), http://kirkmcd.princeton.edu/examples/statmech/szilard\_zp\_53\_840\_29.pdf
  On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings, Behav. Sci. 9, 301 (1964), http://kirkmcd.princeton.edu/examples/statmech/szilard\_zp\_53\_840\_29\_english2.pdf
- [23] L. Brillouin, The Negentropy Principle of Information, J. Appl. Phys. 24, 1152 (1953), http://kirkmcd.princeton.edu/examples/statmech/brillouin\_jap\_24\_1152\_53
- [24] C.H. Bennett, Demons, Engines and the Second Law, Sci. Am. 257(5), 108 (1972), http://kirkmcd.princeton.edu/examples/statmech/bennett\_sa\_257-5\_108\_72.pdf
- [25] J. Earman and J.D. Norton, Exorcist XIV: The Wrath of Maxwell's Demon. Part I. From Maxwell to Szilard, Stud. Hist. Phil. Mod. Phys. 29, 435 (1998). http://kirkmcd.princeton.edu/examples/statmech/earman\_shpmp\_29\_435\_98.pdf
- [26] J. Earman and J.D. Norton, Exorcist XIV: The Wrath of Maxwell's Demon. Part II. From Szilard to Landauer and Beyond, Stud. Hist. Phil. Mod. Phys. 30, 1 (1999). http://kirkmcd.princeton.edu/examples/statmech/earman\_shpmp\_30\_1\_99.pdf
- [27] C.H. Bennett, Notes on Landauer's principle, reversible computation, and Maxwell's Demon, Stud. Hist. Phil. Mod. Phys. 34, 501 (2003).
   http://kirkmcd.princeton.edu/examples/statmech/bennett\_shpmp\_34\_501\_03.pdf
- [28] Z.K. Silagadze, Maxwell's Demon through the Looking Glass, Acta Phys. Polon. B 38, 101 (2007). http://kirkmcd.princeton.edu/examples/statmech/silagadze\_acp\_b38\_101\_07.pdf
- [29] K. Maruyama, F. Nori and V. Vedral, The physics of Maxwell's demon and information, Rev. Mod. Phys. 81, 1 (2009).
   http://kirkmcd.princeton.edu/examples/statmech/maruyama\_rmp\_81\_1\_09.pdf
- [30] X. Oriols and H. Nikolić, Three types of Landauer's erasure principle: a microscopic view, Eur. Phys. J/ Plus 138, 250 (2023). http://kirkmcd.princeton.edu/examples/statmech/oriols\_epjp\_138\_250\_23.pdf
- [31] J.D. Norton, The Simply Uninformed Thermodynamics of Erasure, (Sept. 24, 2023). https://sites.pitt.edu/~jdnorton/papers/Erasure.pdf



