4

Demonic Heat Engines

Benjamin W. Schumacher
Department of Physics
Kenyon College
Gambier, OH 43022, USA

4.1 Demons and the Second Law

One of the many equivalent formulations of the Second Law of Thermodynamics is a limit on the efficiency of heat engines (Fermi 1936). A heat engine does work by exploiting a temperature difference between two parts of its environment; according to the Second Law, its efficiency η is bounded by

$$\eta \leq 1 - \frac{T_2}{T_1} \tag{4.1}$$

where T_1 and T_2 are the high and low working temperatures, respectively. The maximum efficiency is recognizable as the efficiency of a Carnot engine, in which an ideal gas is employed as a working fluid in a reversible cycle of adiabatic and isothermal processes.

Recently, new insight into the Second Law has been obtained by re-examining a famous thought-experiment, Maxwell's demon (Maxwell 1871). The demon is a microscopic being that observes the motions of individual molecules in an equilibrium system and then uses this information to do work, thus violating the Second Law. An important simple example of this idea was proposed by Szilard (Szilard 1929). Szilard's demon operates on a one-molecule gas confined in a chamber. The demon first locates the molecule, then introduces a moveable partition into the chamber. Since the demon knows which side of the partition the molecule is on, it may expand the one-molecule gas by moving the partition, obtaining work.

Szilard claimed the quantum-mechanical process of measurement by which the demon observes the gas molecule will require the demon to expend an energy $\bar{k}_B T$, just offsetting the work obtained. (The shorthand $\bar{k}_B$ represents $k_B \ln 2$ throughout my discussion.) In other words, Szilard supposed that the acquisition of information in a quantum measurement necessarily involves dissipation, and that this dissipation is just sufficient to rescue the Second Law.

More recently, Bennett (Bennett 1982) and Zurek (Zurek 1989) have both pointed out that the *acquisition* of information is not necessarily irreversible, even in quantum mechanics. They view Maxwell's demon as a kind of computer, which acquires and processes information during its operation. Bennett has shown that computers may function in a completely reversible way, and that the only computer operation that requires energy dissipation is the *erasure* of information, which costs at least $\hbar_B T$ per erased bit. According to Bennett, therefore, the demon fails because it must *erase* information to operate in a closed cycle, and the erasure of N bits of information will dissipate an energy $N\hbar_B T$. Zurek points out that, even if the demon retains its stored information without erasing anything, the Second Law may be salvaged by including an additional term in the entropy reflecting the *algorithmic randomness* of the demon's memory state. On average, each bit in the computer's memory adds $\hbar_B$ to the total entropy.

4.2 A Demonic Engine

Maxwell's demon fails in its task because it operates in an environment that is in equilibrium. The energy *value* of acquiring information is equal to the energy *cost* of free memory in which to store the information.

On the other hand, suppose the demon is placed in an environment that is not in equilibrium? It should then be possible for a demon to accomplish work by extracting energy from its surroundings. For example, we might consider a demon placed between two thermal reservoirs of different temperature. The energy value of information (about the warm reservoir) would then exceed the energy cost of free memory (relative to the cool reservoir). How efficiently can the demon function as a heat engine?

Let us begin by considering a version of Maxwell's demon which operates in a completely reversible way, and which we might call a "standard" demon. It consists of three parts: a one-molecule Szilard gas, a reversible computer, and a reversible energy storage device (or "battery"). These subsystems are coupled to each other and to the environment in the following ways:

- The Szilard gas can be thermally coupled to an external thermal reservoir.
- The Szilard gas is mechanically coupled to the battery, so that work performed by the gas may be reversibly stored in the battery and energy from the battery may be used to do work upon the gas.
- Memory elements in the computer may be reversibly coupled to the Szilard gas. The volume of the gas cell may be divided into two equal halves, which we will denote A and B. The molecule may be in either half. A particular memory element has exactly two states, 0 and 1. The coupling connects the location of the molecule with the state of the memory element in the following way:

$$\begin{array}{ccc} A\,0 & \longrightarrow & A\,0 \\ A\,1 & \longrightarrow & B\,0 \\ B\,0 & \longrightarrow & B\,1 \\ B\,1 & \longrightarrow & A\,1 \end{array}$$

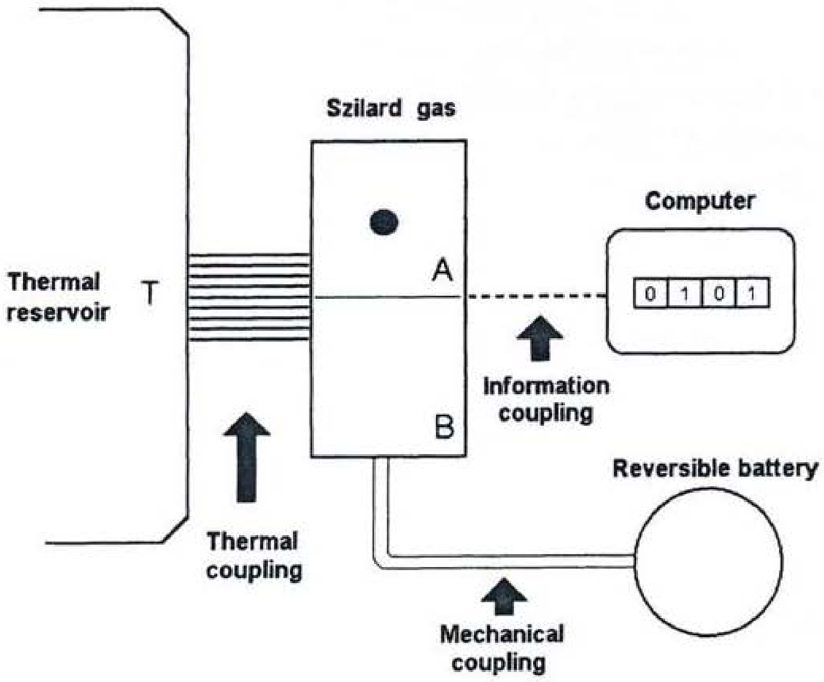

Fig. 4.1. The standard demon.

Note that if the memory element is initially 0 (blank), this coupling causes the memory element to record the location of the molecule in the code 0 = A, 1 = B.

We may also imagine that the demon may exchange energy from the battery and information from the computer with other systems. A schematic of the standard demon is shown in Figure 4.1.

A standard demon connected to a thermal reservoir at a temperature T has two operating cycles, called READ and ERASE. In the READ cycle, the demon first locates the molecule in the Szilard gas and then allows the gas to expand isothermally. This will increase the amount of information in the computer's memory by one bit and add a mean energy of $k_B T$ to the battery. In the ERASE cycle, the demon first compresses the Szilard gas isothermally into region A, then transfers the contents of one bit of its memory to the gas using the information coupling. Note that the coupling leaves the demon's memory in the state 0, so that the bit has been erased; however, the initial compression of the gas requires an average energy $k_B T$ from the battery.

Clearly, a standard demon operating in a READ/ERASE double cycle connected to a thermal reservoir can gain no net energy on average, since the ERASE cycle costs as much energy as the READ cycle provides. On the other hand, suppose our system included *two* demons connected to a different thermal reservoirs and able to exchange information and energy. The information acquired in the READ cycle of one demon may now be disposed of in the ERASE cycle of the other demon, perhaps at a reduced cost. Leftover energy could be used to lift weights, etc.

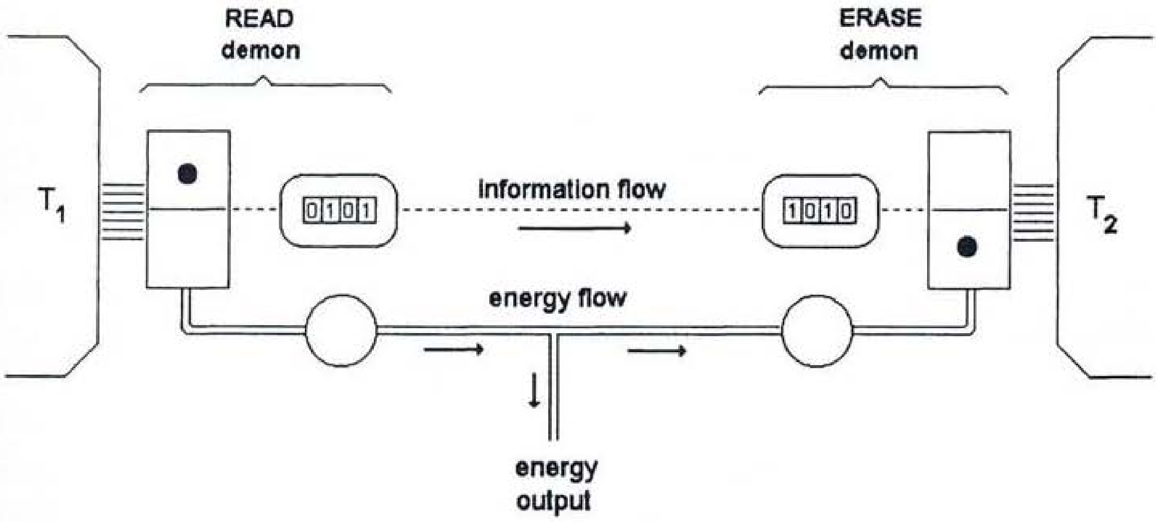

Fig. 4.2. A demonic heat engine.

A diagram of this arrangement is shown in Figure 4.2. This pair of demons functions as a *demonic heat engine*, exploiting the temperature difference between the reservoirs to do work by acquiring and erasing information. In the demonic heat engine, the READ cycle is performed on the thermal reservoir at the higher temperature T_1; then information and mechanical energy is reversibly transfered to the second demon, which performs the ERASE cycle on the cooler reservoir at temperature T_2.

The demonic heat engine has aquired an energy of $k_B T_1$ in each cycle, but it has been forced to use some of this energy ($k_B T_2$) to dispose of unwanted information. It is thus easy to see that the demonic heat engine has a thermodynamic efficiency

$$\eta = 1 - \frac{T_2}{T_1}$$

which is, of course, simply the efficiency of an ideal ordinary heat engine. The READ/ERASE cycle is thermodynamically equivalent to the Carnot cycle.

4.3 Remarks

In an ordinary heat engine, it is easy to identify the heat flow between the two thermal reservoirs. In the demonic heat engine, however, there are separate transfers of information and mechanical energy. This indeed constitutes a flow of heat provided we adopt Zurek's suggestion and count the algorithmic entropy of the transfered bits as "entropy flow".

The demonic heat engine makes use of a temperature difference between two reservoirs; but other demonic engines could make use of other non-equilibrium degrees of freedom in their environments. Suppose, for example, that some order in the thermal reservoir caused the measured positions for a standard demon's Szilard gas to follow a highly correlated sequence: AAAAAABBBBBBAAAABBBBB ....

In other words, the effect of the environment on the demon's "sensing" apparatus is partially predictable, since A is usually followed by A. The demon could easily use this order to reversibly *compress* the information acquired in its READ cycles; then it would need to run fewer ERASE cycles to clear its memory. Even though each ERASE cycle might cost as much energy ($k_B T$) as each READ cycle provided, the demon could still succeed in acquiring free energy from its environment.

In short, the predictability of the environment constitutes an effective temperature difference between the READ and ERASE cycles of the demon. An environment will present two effective temperatures to a demonic engine: a READ temperature T_R, which measures the energy value of acquired information, and an ERASE temperature T_E, which measures the energy cost of free memory. In general, $T_R \geq T_E$, with equality at thermal equilibrium. The efficiency of the demonic engine will be

$$\eta \leq 1 - \frac{T_E}{T_R}. \tag{4.2}$$

In a sense, we ourselves are demonic engines. We gain access to energy through the information we possess about our environment. The effective READ temperature of our environment is astronomically high: even the sum of the genetic, cultural, and sensory information available to human beings is small compared to thermodynamic entropies, yet the free energy that this makes available to human beings is enormous. The ERASE temperature is a few hundred Kelvins, but free energy is so easily available that we typically store and process information in thermodynamically "wasteful" ways.

I have one final observation that is more directly germane to the subject of this workshop. In the demonic heat engine, the demon at the warm reservoir transfered the contents of its memory to the demon at the cool reservoir, clearing its own memory cells. Suppose instead that the first demon sent over a *copy* of its information, retaining its own record afterward. The ERASE cycles of the cool reservoir demon would not now restore the demonic heat engine to its initial state, and more erasure would be required. In other words, for the demonic heat engine to function at its ideal efficiency, the efficiency at which the total entropy of the entire system remains constant over time, *no copy can be made of the transfered information*.

So the *increase* in entropy in a demonic heat engine is related to the *copying* of information. If information is copied within the demon—or indeed, if an external observer copies information from the demon's memory by examining it—entropy increases during the operation of the demonic heat engine.

Since this talk comes near to the beginning of the workshop, let me venture a modest prediction. I predict that most of the suggestions made here about the physical origin of time asymmetry will boil down to Nature's enormous propensity for *making copies* of information. Why Nature should be so constituted is a deeper mystery, and a much harder problem.

Note added after the workshop: I am not certain to what degree my prediction has been fulfilled in the discussions of decoherence and quantum cosmology that have dominated much of our agenda. It is arguable that the "leaking away" of quantum phase relations into the external environment that is so crucial in models of decoherence is an instance of the uncontrolled copying of information in Nature.

Discussion

Davies You are using the concept of temperature to describe a non-equilibrium system. What exactly do you mean?

Schumacher I mean nothing more than the ratio between the energy transfer and entropy change of a system: $\Delta E = T\Delta S$. To be careful, we should call this an "effective temperature".

Bennett (1) Szilard almost understood it. At the end of his paper, he did a calculation in which entropy only increased during erasure. (2) The representation of the demon you use [on your transparencies] is that used by Larry Gronik in his cartoon about Maxwell's Demon in the July 91 issue of Discover magazine.

Schumacher You are quite right about Szilard. As for my cartoon demon, I got it from the figure on the Underwood deviled ham can.

Gell-Mann So-called "creation scientists" sometimes argue that biological evolution is impossible because it allegedly violates the second law of thermodynamics. Some real scientists occasionally claim that biological evolution and related phenomena are instances of local exceptions to the second law.

If we take into account the environment, with its regularity, and the algorithmic complexity term in the entropy (where it is appropriate), we see that evolution that adapts, more or less, to those environmental regularities represents an example of the second law and not an exception.

Schumacher Yes, I agree.

Lloyd Two points: First, the amounts of information to which we are accustomed are negligible compared to normal thermodynamical quantities. All the libraries in the world contain much less information than that in a gallon of gas.

Second, our normal heat engines are already "demonic": their ability to do work depends crucially on their ability to get and process information.

Schumacher This is a very good point. All heat engines must process information. This is true even of the Carnot engine: it must know which of the heat reservoirs is hot and which is cold. Otherwise it might act as a refrigerator!

Davies Following from Seth Lloyd's remarks: Is the fact that a copy of Encyclopaedia Britannica costs much more than a gallon of gasoline a reflection of the fact that the information content of the former has greater "value" (or depth) than the latter?

Schumacher Perhaps the encyclopaedia is valuable because the information is strongly correlated to the rest of the universe. It contains information about the composition of air, how to obtain gasoline, how to build an internal combustion engine, etc. One could imagine a very clever demonic engine which read the encyclopaedia, then went off and used the information to obtain lots and lots of energy. This would correspond in my scheme to an extremely high READ temperature—aquisition of a relatively small amount of information carries with it access to a huge amount of energy.

Omnès Would you please elaborate upon what you said last, i.e., nature is constantly making copies of itself.

Schumacher I have in mind, among other things, the phenomenon of decoherence, about which so much is said at this conference. Information about quantum correlations gets copied into many environmental degrees of freedom, beyond any hope of recovery, and so phase relationships between branches of the wave functions are effectively destroyed. But analogous things happen in classical systems as well.

Hartle Does the measure of irreversibility supplied by the "multiplication of copies" agree quantitatively with entropy augmented by complexity, or disagree with it?

Schumacher I am not sure. I have not really proposed a quantitative measure of irreversibility based on copying; however, that is a very nice idea. I would be surprised if it did not work out.

Zeh Is the fact that "Nature likes to make copies" not essentially equivalent to the retardation of radiation (which carries away information)? Philosophers like to call it the "fork of causality".

Schumacher Yes, certainly they are related. I would rather say that the retardation of radiation is a crucial example of Nature's propensity for copying information.

Albrecht Is not the ability of nature to make all these copies in turn related to the out-of-equilibrium state of the universe? This is reflected in the presence of heat baths at two different temperatures in the apparatus you describe, or, for example, in the availability of blank tape which a demon could put to good use.

Schumacher Yes, I think that you are right. If you try to copy something onto a tape that already contains random bits (using some reversible writing scheme, such as the XOR function), you just change the random bits to another set of random bits. You have not made the tape any harder to erase. If copying of information does occur, there must be lots of "blank tape" in the universe—that is, the universe must be far from equilibrium.

Starobinsky If you really want to develop further your proposal that the source of irreversibility in nature lies in unconscious multiple copying of information, you have to take into account unavoidable errors that appear in this process. As a result, only approximate copies will be produced which can equally well be considered as containing new information. To check if they are identical to an original and to correct errors (if desired) will cost more energy and will produce more entropy. We all encounter this problem when typing our papers or sending them by e-mail.

Schumacher Errors in copying will, of course, only make things worse for erasing, and so contribute to the irreversibility of the process. In this sense, copying errors would be a source of "friction" in a demonic heat engine, reducing the actual efficiency still further from the ideal efficiency.

References

Bennett, C. H. (1982) The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, **21**, 905–940.

Fermi, E. (1936) *Thermodynamics*, Dover, New York.

Maxwell, J. C. (1871) *Theory of Heat*, Longmans, London.

Szilard, L. (1929) On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. (In German) *Zeitschrift für Physik*, **53**, 840–856. Translation reprinted in Wheeler, J. A. and Zurek, W. H., eds., (1983) *Quantum Theory and Measurement*, Princeton University Press, Princeton.

Zurek, W. H. (1989) Algorithmic randomness and physical entropy I. *Physical Review* A **40**, 4731–4751.