When is information erasure necessary between successive thermodynamic cycles ?

François Barriquand
Maison paroissiale, Place de l'église,
94320 Thiais, France

Abstract

A modified version of Leo Szilard's 1929 thought-experiment is considered, which allows one to show that the iteration of some specially conceived thermodynamic “cycles” does not require any costly memory erasure. These thermodynamic cycles can be easily examined from either a classical or a quantum point of view. Although they differ randomly from each other in certain aspects, their remarkable energy balance remains systematically constant. Similar cycles can prove useful for lowering the theoretical cost of memory erasure in computers.

Introduction

In 1929, Leo Szilard[1] proposed to use a single-molecule thermodynamic “fluid” in order to perform an elementary thought-experiment which challenged the understanding of physicists during many years. After some hesitation, one came to realize that quantum measurements do not necessarily need to be thermodynamically costly. This conclusion rendered Szilard's model even more intriguing. In 1961, Rolf Landauer[2] pointed out that the cost of the *erasure* of information between thermodynamic cycles should be carefully taken into account, even when information *acquisition* appears to be costless. In 1970, Oliver Penrose[3] developed a similar idea, which seemed to reconcile Szilard's thought-experiment with the second principle of thermodynamics.

Szilard's model, which exhibits a kind of “binary” structure, has also served as a basis to explore the minimal thermodynamic costs of computation, which must include the preparation steps that precede computation itself. In particular, if one wishes to perform several computations with the same computer, each of them using the same quality and quantity of memory, one needs to *reset* the *memory* by *erasing* it before each new computation. In the preceding sentence, the three terms “reset”, “memory” and “erasing” all belong to the semantic field of computer science. Their corresponding meanings can be translated into physics, which allows one to estimate the thermodynamic cost required for erasing one “bit” of memory. A rigorous calculation of erasure costs has been performed by Barbara Piechocinska[4], among others. Within the framework of her study, elementary erasure costs always amount to at least k_BTLog2, which precisely corresponds to the amount needed to save the second law, as R. Landauer had indicated.

In spite of this seemingly reassuring conclusion, nobody has ever formally proved that “erasure” should be absolutely necessary for all kinds of iterated procedures analogous to Szilard's experiment. What is more, in spite of B. Piechocinska's calculations, nobody, either, has ever proved that the minimal cost of erasure of one “bit” of data should always amount to k_BTLog2 within *any* physical framework. As Karl Popper[5] has shown, the general character of physical laws always remains, at best, conjectural, since the discovery of a single counter-example suffices to cancel the universal validity of a theory. As far as Szilard's model is concerned, one may wonder whether borrowing the concept of memory “erasure” from computer science has not led physicists to analyze Szilard's experiment too narrowly, consciously or not. The first section of my present study suggests that this is precisely what has happened. In this section, I exhibit a variant of Szilard's scheme for which no costly memory erasure is necessary. A subsequent reevaluation of minimal “memory erasure” energy costs is proposed in section B. The question of how to maximize, from a fully macroscopic perspective, the tiny energy gain considered in section A is raised in section C.

A. Revisiting Szilard's thought-experiment : can the need for quantum erasure be circumvented ?

Let us first follow Szilard's original proposal by imagining a hollow cylinder that can be separated in two halves by a partition of adjustable height. From the point of view of classical physics, adjusting the height of this partition (*i.e.* closing or opening it) can be achieved without modifying the energy of any molecule contained in the cylinder, provided that the partition is thin enough. From the point of view of quantum physics, the situation appears less simple, since modifying the boundaries of a quantum wave function is usually not energetically innocuous. Quantitatively speaking, however, quantum studies, for instance by Wojciech H. Zurek[6], have indicated that adjusting the partition's height does not cause any significant change in the energy of a one-molecule quantum fluid, at least as long as the temperature of the fluid is high enough. This result ensures the continuity between a fully quantum and a classical description of Szilard's machine.

According to Szilard's original proposal, the partition, which is initially open, must be closed at the beginning of a new cycle ($t=t_0$) by the machinery used for the experiment. At a later time $t=t_1$, the “position” (whether in the left or the right half of the cylinder) of the one-molecule fluid is measured. This information is recorded with one "bit" of memory. At time $t=t_2$, a piston is placed on the side opposite to the one-molecule fluid, against the partition. The partition is opened at time $t=t_3$. Between $t=t_3$ and $t=t_4$, the one-molecule fluid is decompressed isothermally, which allows the operator to store up an amount of work equal to $k_B T \mathrm{Log} 2$. At time $t=t_5$, the partition is closed again. At the same time, the one-bit memory is "erased", which costs an amount of work equal to... $k_B T \mathrm{Log} 2$. A new cycle can now begin. None of the above-described successive cycle(s) can violate the second law of thermodynamics, but they are also quite useless.

Let us now consider an even simpler machine, whose piston can always remain on the right side of the one-molecule fluid. My machine still requires the use of a one-bit memory, which can take the values "0" or "1", as in Szilard's scheme. The details of the modified cycle I now present are the following :

- Just before the beginning of the cycle, the one-bit memory is in state "0". The one-molecule fluid can move freely within the entire cylinder, whose partition is opened. At time $t=t_0$, the partition is closed. From a quantum mechanical point of view, if both sides of the partition remain perfectly rigid and non-absorbing, the wave function of the molecule can still "occupy" both sides of the cylinder simultaneously with equal probability. Decoherence, as described below, can remove this ambiguity starting from time $t=t_1$.

- In order to ensure that the one-molecule fluid is localized on a single side of the cylinder at time $t=t_1$, let us suppose that the corresponding molecule possesses a spin. Efficient localization of the molecule can be achieved *via* two macroscopic spin reservoirs, one coupled to the one-molecule fluid when it is located on the left side of the cylinder, the other when it is on the right side. In fact, a single spin reservoir, located either on the left or the right side of the cylinder, interacting with the spin of the one-molecule fluid *via* spin operators S_z oriented along a single identical axis, would suffice to localize the particle, as illustrated in the textbook of Maximilian Schlosshauer with the help of a well-known model which he calls "simple model for decoherence"[7]. The "simple model" presented by M. Schlosshauer is exactly solvable. It guarantees us that decoherence can be both efficient and energetically virtually costless, provided that the time allowed for it is long enough. If we further imagine that our spin reservoir interacts with a thermostat, we can also be assured that this reservoir can be costlessly "reset" (*i.e.* : reach a new random starting point) before another cycle begins, which allows us to use the same spin reservoir for an unlimited number of cycles.

- At time $t=t_2$, the one-molecule fluid interacts with our one-bit memory during an adequately determined duration, so as to update the information contained in the memory. This can be done costlessly with the help of the following Hamiltonian, wherein W represents a constant which can be chosen as small as one wishes :

$$\mathcal{H} \; = \; W.\ |\textit{single-molecule fluid on the left}\rangle\langle\textit{single-molecule fluid on the left}|\otimes$$
$$\{|\textit{memory in state ``1''}\rangle\langle\textit{memory in state ``0''}| + |\textit{memory in state ``0''}\rangle\langle\textit{memory in state ``1''}|\}$$

If the single-molecule fluid is located within the right half of the cylinder, $\mathcal{H}$ does not modify anything : the memory status does not change. On the contrary, if the one-molecule fluid is located within the left half of the cylinder; $\mathcal{H}$ modifies the memory state from “0” to “1”.

- At time $t=t_3$, if the memory is still in state “0”, the partition is opened, then closed again. The machine then repeats exactly what it has already done at $t=t_1$ (localization of the fluid) and $t=t_2$ (memory update). If the memory is still in state “0” after this, the partition is opened again, then closed... the machine again repeats what it has done at $t=t_1$ (localization of the fluid) and $t=t_2$ (memory update), *etc.* This iteration ensures us that at a certain time $t=t_n$, the one-molecule fluid will eventually find itself within the left side of the cylinder, so that the memory state will change from “0” to “1”. It is interesting to note that the value of time t_n cannot be known in advance, so that different “cycles” are not strictly equivalent in all aspects. From a thermodynamic point of view, however, it is still legitimate to call them “cycles”, since their energy balance remains strictly the same.

- Let us now place ourselves at time $t=t_n$. As stated by Charles H. Bennett, no microscopic quantity can correspond to entropy, “frustrating the natural desire of molecular model builders to regard an individual molecular configuration as having an entropy”[8]. In contrast, it must be emphasized that temperature itself, quoting the words of C. H. Bennett, “can be expressed as the average of $mv^2/2$”, which allows one to consider it as a statistically meaningful quantity, even for a single molecule. From a classical point of view, the situation of our system at $t=t_n$ appears therefore quite simple to describe : the temperature of the one-molecule fluid has remained unchanged; its pressure has been doubled ; the volume within which it evolves has been halved. This allows us to use our piston in the same way as Szilard would have used it, according to the following steps : (i) place the piston against the right side of the partition ; (ii) open the partition ; (iii) move isothermally the piston towards the right side of the cylinder, in order to store up a quantity of work equal to k_BTLog2.

- Once all the preceding steps have been performed, our one-bit memory necessarily finds itself in state “1”. In order to prepare a new cycle, we need to modify this state from “1” to “0”, which can be done at a suitable time $t=t_{n+1}$. What is absolutely crucial for us here is that such a modification can be performed without any costly energy expenditure. If we needed to *erase* a random memory configuration such as “either 0 or 1”, as demanded in Szilard's original thought-experiment, finding a way to spend less than k_BTLog2 of energy would be quite challenging (although I shall show in section B below that this is indeed possible). But here, in my own version of Szilard’s experiment, I only need to move the one-bit memory from the well-known state “1” to the well-known state “0”, which can be done at an arbitrarily low cost. In other words, although some kind of memory *resetting* between successive cycles still remains necessary within my new scheme, this scheme allows one to circumvent the need for any costly memory *erasure*.

Another “cycle” can now begin. Each cycle allows one to convert the same quantity of heat, k_BTLog2, into an equivalent quantity of work. As a consequence, the second law of thermodynamics does not hold.

Where lies the magic ? In a way, we may consider that the one-molecule fluid itself has served as a temporary memory, whose past can be erased costlessly by decoherence during each cycle. The entire procedure seems so simple that I have not needed to challenge the prevalent opinion, according to which the minimal cost of erasing one-bit of memory amounts to k_BTLog2. But does this opinion still hold ? Let us now re-examine this question, in the light of my modified use of Szilard's engine.

B. What is the minimal cost of quantum erasure ?

Can a procedure analogous to the one described above also enable us to decrease the minimal cost of memory erasure ? As we shall soon see, the answer to this question is *yes*.

My basic idea is quite simple : it consists in using Szilard's machine itself as memory. We may decide, by convention, that when the one-molecule fluid is confined within the left (resp. right) portion of the cylinder, Szilard's machine is in state “0” (resp. “1”).

Let us now start from an initially random memory state equal to “either 0 or 1”, which we wish to transform into the well-known “0” state in the most economical way. In order to simplify my presentation, I shall avoid the indexation of successive physical instants $t=t_1$, t_2, t_n..., drafting my entire procedure in the style of a computer program :

[*step* 1]: – Open the partition, in order to delocalize the one-molecule fluid.

[*step* 2]: – Close the partition and localize the fluid, either on the left or the right side of the cylinder, *via* the decoherence process induced by a macroscopic spin reservoir, as described in section A above.

[*step* 3]: – Approach a special apparatus *A* towards the left part of the cylinder. The interaction of *A* with the one-molecule fluid can be described by means of the following Hamiltonian (whose characteristics are quite similar to those of the Hamiltonian considered in section *A* above) :

$$\mathcal{H} \;=\; W.\ |\textit{fluid on the left}\rangle\langle\textit{fluid on the left}| \otimes$$
$$\{|\textit{different state of }A\rangle\langle\textit{initial state of }A| + |\textit{initial state of }A\rangle\langle\textit{different state of }A|\ \}$$

[*step* 4]: – If *A* finds itself in a different state than its initial state, go to [*step* 5].
– If, on the contrary, *A* remains in its initial state, go to [*step* 1].

[*step* 5]: – Return *A* to its initial state.

As can be easily seen, when the entire procedure reaches the end of *step* 5, Szilard's machine finds itself in the well-known state “0”. No energy has been spent by the operator.

As a consequence of Landauer's analysis, this result provides a supplementary proof that the second “law” fails to be universally valid, as already shown in section A above.

C. Macroscopic optimization.

The procedure described in section A above only allows an operator to gain a quantity of work equal to k_BTLog2 per cycle, which, from a macroscopic point of view, is a nearly insignificant quantity. Is it possible to do better ?

In theory, the answer is obviously affirmative. The simplest way to store up more energy would consist in using a greater number *N* of identical Szilard's engines, which would allow one to multiply by *N* the gains obtained during one cycle. From a practical point of view, however, building up *N* identical machines can hardly be considered as a meaningful improvement.

Trying to increase thermodynamic gains with a *single* machine represents a much more difficult challenge. Naturally, Szilard”s machine could be made more “macroscopic” by using a fluid composed of *N* molecules instead of one only ; unfortunately, separating a *N*-molecule fluid in two halves with a partition would only generate an average pressure difference between both sides proportional to $\surd N$. With *N* molecules, all sorts of undesired thermodynamic noises would also increase within the same proportion $\surd N$, which seems rather discouraging.

From a more positive angle, it is worth noting that I have already proposed a quite different thought-experiment in 2005[9], which, as I have claimed, can also break the second law. In contrast with Szilard's model, my 2005 scheme does not use any specific “memory” at all. What is more, its cycles can all be accomplished within the same duration. In spite of these huge discrepancies, it is rather likely that a hidden continuity may be found between both of my models. In particular, it is worth noting that, in spite of its unusual characteristics, Szilard's model can function quite smoothly within a classical perspective, as is also the case for my 2005 scheme. None of them relies in a very crucial way on any subtle quantum phenomena such as decoherence or “measurement”. Although the purpose of my present article is not to examine my 2005 model, one may estimate that the mere possibility that a connection may exist between two markedly different schemes increases the chances that a rather wide continuous range of connate strategies may eventually prove successful in breaking the second law. From the experimental point of view, this perspective looks rather comforting. I may also add that the general ideas lying behind my 2005 scheme could possibly serve as a more convenient source of inspiration for experimentalists than Szilard's machine, since the partition used within Szilard's machine seems particularly inconvenient to miniaturize.

In spite of such optimistic considerations, it must also be recognized that neither my 2005 scheme nor my 2014 scheme offer easily affordable ways to convert substantial, macroscopic amounts of heat into work. This practical difficulty might lie behind the fact that *no* living organisms, such as bacteria, seem to have ever developed efficient procedures to convert heat into work ; chemical sources of energy, plus sunlight, offer them a large panel of more convenient and more efficient opportunities.

From a purely statistical perspective, it is instructive to bear in mind that if it were algorithmically possible to compress the memory size needed to store randomly measured numbers, the second law could be invalidated independently from the thermodynamic details of any microscopic experiment. In other words, if the second law of thermodynamics were universally true, physical arguments would never suffice to justify it without the help of number theory. This being said, it must be added that series of random physical measurements “always” (statistically speaking) provide results which happen to be incompressible, as shown by Gregory Chaitin[10]. The second law of thermodynamics would therefore remain quite safe if it remained in the hands of statisticians and computer programmers alone. Even too safe... since certain physical procedures can serve to invalidate the so-called second “law”, as I have shown.

Conclusion

The present article relies on a model of great simplicity, which can be easily analyzed from both a classical and a fully quantum point of view. By showing that Landauer's assumption concerning the ultimate necessity and cost of memory erasure is not universally valid, I have demonstrated that the second law itself is not always reliable, as I had already shown in 2005. I hope that the simplicity of the present article will eventually convince some readers that both my present claims and those contained in my 2005 article devoted to thermodynamics are valid.

This being said, it would be quite unreasonable to hope that converting microscopic or, at best, extremely tiny quantities of heat into work could have any short-term impact on humanity's energy consumption. For the time being, moderation in the use of fossil sources of energies, and the promotion of sustainable lifestyles, figure among the simplest methods that might serve to limit the ecological damage suffered by our planet.

References

[1] L. Szilard, *On the decrease of entropy in a system by the intervention of intelligent beings*, Z. f. Physik **53**, 840–856 (1929).

[2] R. Landauer, *Irreversibility and Heat Generation in the Computing Process*, IBM J. Res. Dev. **5**, 183–191 (1961).

[3] O. Penrose, *Foundations of Statistical Mechanics*, Pergamon Press, Oxford (1970).

[4] B. Piechocinska, *Information Erasure*, Phys. Rev. A **61**, 1–9 (2000).

[5] K. Popper, *Logik der Forschung*, Springer (1935).

[6] W. H. Zurek, *Maxwell's demon, Szilard's engine and quantum measurements*, in G. T. Moore and M. O. Scully, *Frontiers of Nonequilibrium Statistical Physics*, Plenum Press, New York, 151–161 (1984).

[7] M. Schlosshauer, *Decoherence and the quantum-to-classical transition*, chap.2, pp. 88–93, Springer (2007).

[8] C. H. Bennett, *The Thermodynamics of Computation – A Review*, Int. J. Theor. Phys. **21**, p. 314 (1982).

[9] F. Barriquand, *Perte d'information et irréversibilité en thermodynamique*, Revue des Questions Scientifiques, **176** (1), 59–74 (2005). An English version of this article, without figures, is available at http://arxiv.org/abs/physics/0507184

[10] G. Chaitin, *Meta Math ! The Quest for Omega*, Vintage Books, 2005.