2-20-2023

A mathematical investigation of Landauer's principle

Paul Bracken
The University of Texas Rio Grande Valley, paul.bracken@utrgv.edu

Recommended Citation

Bracken, P. A mathematical investigation of Landauer's principle. J Math Chem 61, 1292–1300 (2023). https://doi.org/10.1007/s10910-023-01463-x

A Mathematical Investigation of Landauer's Principle

Paul Bracken

Department of Mathematics,

University of Texas,

Edinburg, TX, USA 78540

paul.bracken@utrgv.edu

Abstract

A minimal mathematical approach is used to state Landauer's principle in a precise, general way. The results are obtained by means of a rigorous development which is based on the use of quantum statistical mechanics. A mathematical form of the principle results as an equality rather than an inequality. The equality version does imply the original statement of the principle as introduced by Landauer.

PACS: 42.60.Da

Keywords: probability, density matrix, reservoir, entropy temperature, erasure, second law

1 Introduction

The paradox referred to as Maxwell's demon first appeared in the nineteenth century as the study of thermodynamics developed and quantum mechanics had not yet been developed [**1-3**]. The traditional way of presenting the paradox was to say the entropy of a gas of particles can be lowered without expending energy and thus violating the second law of thermodynamics. This statement overlooked the fact that there is information about the positions and momenta of the particles which has accumulated in the demons memory. For a closed cycle, this must be taken into consideration. The demons memory must also be returned to its initial state to properly close a cycle of the system which means there is erasure of information. Erasure of information always produces heat, even when only one bit is erased. Thus thermodynamics by itself places physical constraints on information processing. In fact both quantum mechanics and thermodynamics should be used together in studying the paradox [**4**]. Landauer realized that it is irreversible erasure of information that logically demands a corresponding entropy increase in the environment. Information erasure with respect to the degrees of freedom which carry information causes entropy to flow to those degrees of freedom which do not carry information. At inverse temperature β, this entropy increase causes the dissipation of heat and satisfies an inequality

$$\Delta Q \geq \Delta S / \beta. \tag{1.1}$$

In (**??**), ΔS is the entropy decrease in the memory. Proposal (**??**) is now referred to as Landauer's principle, and (**??**) is known as the Landauer bound or limit.

Instead of a classical probability distribution, at the quantum level, a bit is described by means of a density matrix ρ and the corresponding relevant entropy is the von Neumann entropy. With this concept provided by quantum mechanics, a quantum version of the second law can be derived. The quantum version makes the same prediction as its classical counterpart. For the case in which a maximal entropy quantum bit is completely erased, this quantity of heat is $k_B T \log(2)$, as expected considering (**??**).

It is the intention here to formulate the exact setting of Landauer's principle in precise mathematical and statistical mechanical terms. Also we investigate how (**??**) can be obtained in a mathematically precise, rigorous way [**5-7**]. It can be formulated to include various processes which are more general than erasure. The setting is minimal in the sense that the Landauer bound can be violated when any one of the assumptions is discarded. Suppose $\mathcal{S}$ denotes system and $\mathcal{R}$ reservoir. The main result is a proof based on the axioms of this result in the generalized form

$$\beta \, \Delta \, Q = \Delta \, S + I(\mathcal{S}'; \mathcal{R}') + D(\rho'_{\mathcal{R}} || \rho_{\mathcal{R}}) \geq \Delta \, S. \tag{1.2}$$

The variable $I(\mathcal{S}'; \mathcal{R}')$ in (**??**) is the mutual information between system and reservoir. It quantifies the correlations set up between system and reservoir during the process. The relative entropy $D(\rho'_{\mathcal{R}} || \rho_{\mathcal{R}})$ can physically be interpreted as the free energy increase in the reservoir. Looking at inequality (**??**) more closely shows that the Landauer bound $\beta \, \Delta \, Q \geq \Delta \, S$ can be most precise only when $\Delta \, S = 0$. In fact, the bound (**??**) can be improved for all non-trivial processes.

The explicit improvement of Landauer's bound which is possible when the thermal reservoir which participates in the process has a finite dimensional Hilbert space constitutes a second result. The proof depends to a great degree on the exact system that is realized. The actual setup and subsequent statements will be quantum mechanical in nature. However, they apply to the classical, probabilistic case as well by providing various restrictions. A Hamiltonian for the system would be necessary in order to make other statements such as the work done on the system which the stated bound $\beta \, \Delta \, Q \geq \Delta \, S$ does not do. The total energy of system-reservoir $\mathcal{S}\mathcal{R}$ need not be defined. Even if a Hamiltonian were given for S, the total energy may not be conserved during the process. Finally two physical model examples are discussed to contrast with the formal preceeding mathematical presentation [**8-11**]. It should be emphasized that Landauer's principle does impact the area of quantum thermodynamics, which is of interest to chemistry. A way of expressing it relevant here is if an observer loses information about a system, heat is generated and the observer loses the capacity to extract useful work from that system. Landauer's principle in its general meaning establishes a relation between thermodynamics and logical irreversibility. In nonequilibrium statistical physics, there is no a priori relation between logical and thermodynamic reversibility. It is possible a physical process is logically reversible, but thermodynamically irreversible. Recently, a Landauer erasure was actually performed at cryogenic temperatures (T=1 K) on an array of high spin (S=10) quantum molecular magnets. The array is set to act as a spin register where each non-magnetic encodes a single bit of information. The experiment has set the extension of the validity of the Landauer principle to the quantum realm. On account of fast dynamics and low inertia of single spins used, the experiment showed how an erasure process can be carried out at the lowest possible thermodynamic cost.

2 Physical Picture and Mathematical Formulation

To proceed let us establish the exact mathematical and physical setup for the process. Landauer's process is supposed to erase or reset the state of a system by allowing it to interact with a thermal reservoir. This brings the system to a definite state, such as a fixed pure state.

The process involves a system $\mathcal{S}$ and reservoir $\mathcal{R}$ both of which have a Hilbert space formulation and description. At the initial point, the reservoir $\mathcal{R}$ is initially in a thermal state described in terms of the operator

$$\rho_{\mathcal{R}} = \frac{e^{-\beta \mathbf{H}}}{\mathrm{tr}(e^{-\beta \mathbf{H}})}. \tag{2.1}$$

In case (**??**), $\mathbf{H}$ is the Hamiltonian, a Hermitean operator, and $\beta \in [-\infty, \infty]$ is the inverse temperature. It is the case that initially, the system $\mathcal{S}$ and the reservoir $\mathcal{R}$ are uncorrelated, so the density matrix $\rho_{\mathcal{SR}}$ has the structure $\rho_{\mathcal{SR}} = \rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}}$. The process as a whole is directed by a unitary evolution given by an operator $\mathbf{U}$,

$$\rho'_{\mathcal{SR}} = \mathbf{U}\, \rho_{\mathcal{SR}}\, \mathbf{U}^{\dagger}. \tag{2.2}$$

These hypotheses which yield this kind of setup are reasonable, typical and turn out to be minimal.

There are two subsystems, $\mathcal{S}$ the system and $\mathcal{R}$ the reservoir, on which the process acts. These are modeled as quantum systems each with a Hilbert space of finite dimension, $d_{\mathcal{S}}$ and $d_{\mathcal{R}}$, respectively. A Hamiltonian for the reservoir is given $\mathbf{H} = \mathbf{H}^{\dagger} \in \mathcal{B}(\mathbb{C}^d)$. Suppose the reservoir is in a thermal state (**??**) at inverse temperature β before the process begins. For $\beta = \pm\infty$, $\rho_{\mathcal{R}}$ is the maximally mixed state on the ground state space of $\pm\mathbf{H}$. If $\rho_{\mathcal{R}}$ were not in a thermal state such as (**??**), it would be possible to violate Landauer's bound. This is related to the fact that thermal states are the only completely passive states. This means that from an arbitrary number of state copies, work cannot be extracted by unitary operations alone.

No assumption is placed on the initial state of the system $\rho_{\mathcal{S}}$. It doesn't have to be a thermal state for example. The following results are in fact completely independent of the system Hamiltonian, which need not even be specified. It is required that the system and reservoir be initially uncorrelated

$$\rho_{\mathcal{SR}} = \rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}}. \tag{2.3}$$

This property is important for Landauer's principle to apply. If the initial state $\rho_{\mathcal{SR}}$ were for example such that the reservoir $\mathcal{R}$ had perfect classical correlations with $\mathcal{S}$, a unitary process could reduce the system entropy without heat dissipation or violation of Landauer's bound. The product state (**??**) is standard in the theory of thermodynamics.

It is often assumed that the reservoir $\mathcal{R}$ has not interacted with $\mathcal{S}$. Their states are independent. When system and reservoir have in fact undergone prior interactions, however, they may be correlated. A Landauer bound may be given for this case. As a reservoir which is initially thermal is assumed, correlations in $\mathcal{SR}$ would be unnatural, unless the full initial state $\rho_{\mathcal{SR}}$ were also thermal. However this would require a Hamiltonian interaction term between $\mathcal{S}$ and $\mathcal{R}$.

It is required that the entire process takes place through any unitary transformation evolution

$$\rho'_{\mathcal{SR}} = \mathbf{U} \left(\rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}} \right) \mathbf{U}^{\dagger} = \mathbf{U} \rho_{\mathcal{SR}} \mathbf{U}^{\dagger}, \tag{2.4}$$

where $\mathbf{U} \in \mathcal{B}(\mathbb{C}^d_{\mathcal{S}} \otimes \mathbb{C}^d)$ acts jointly on $\mathcal{S}$ and $\mathcal{R}$. The assumption of unitarity implies that no unspecified environment E participates in the process to take on entropy. This may occur during a dissipative process described by quantum processes. This evolution forces one to explicitly include all resources used during the process in the description by $\mathcal{S}$ and $\mathcal{R}$.

The operator $\mathbf{U}$ may be implemented by a time-dependent Schrödinger evolution, where an interaction Hamiltonian turns on at some instant in time by an external agent, and then is switched off. Consequently, an effective operator $\mathbf{U}$ acts on the initial state $\rho_{\mathcal{SR}}$ and generates final state $\rho'_{\mathcal{SR}}$. The development here does not require any such particular form for $\mathbf{U}$ or dependence on $\mathbf{U}$. It requires just the fact that the process $\rho_{\mathcal{SR}} \to \rho'_{\mathcal{R}}$ was generated by some unitary operator.

The joint final state $\rho'_{\mathcal{SR}}$ of the system and reservoir may end up correlated. Let $\rho'_{\mathcal{S}} = \mathrm{tr}_{\mathcal{R}}(\rho'_{\mathcal{R}})$ denote the final state of the system-state after and $\rho'_{\mathcal{R}} = \mathrm{tr}_{\mathcal{S}}(\rho'_{\mathcal{SR}})$ the final reservoir state. It is not required $\rho'_{\mathcal{S}}$ be a pure state.

3 Landauer's Principle

Some additional mathematical knowledge is required to formulate and discuss Landauer's principle and its consequences. Landauer's principle relates the decrease of entropy of a system defined as

$$\Delta S = S(\rho_{\mathcal{S}}) - S(\rho'_{\mathcal{S}}) \tag{3.1}$$

to the heat transferred to an associated reservoir

$$\Delta Q = \mathrm{tr}\,[\mathbf{H}(\rho'_{\mathcal{R}} - \rho_{\mathcal{R}})] = \mathrm{tr}[\mathbf{H}\rho'_{\mathcal{R}}] - \mathrm{tr}[\mathbf{H}\rho]. \tag{3.2}$$

In (**??**), S denotes von Neumann entropy of the quantum state ρ defined as

$$S(\rho) = -\mathrm{tr}[\rho \log \rho]. \tag{3.3}$$

So (**??**) corresponds to the average increase in internal energy of the thermal reservoir. The use of the term heat is justified as this Initially $\mathcal{R}$ is a thermal reservoir. Thus the term heat may be used as this energy is not in an ordered form. It is possible it may absorb entropy from $\mathcal{S}$ during the process such that energy is allocated over many states. In some of the results, another variable appears, the entropy increase of the reservoir, Δ. This variable which appears in the final result is defined as

$$\Delta = S(\rho'_{\mathcal{R}}) - S(\rho_{\mathcal{R}}). \tag{3.4}$$

There are two other quantities of statistical mechanical interest that come up.

The relative entropy $D(\sigma||\rho)$ between two states σ and ρ is defined to be

$$D(\sigma||\rho) = \mathrm{tr}[\sigma \log \sigma] - \mathrm{tr}[\sigma \log \rho] \tag{3.5}$$

For a state ρ_{AB} of a bipartite system which has reduced states obtained by tracing is given by $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{tr}_A[\rho_{AB}]$. The mutual information is defined as

$$I(A;B) = I(A;B)_{\rho_{AB}} = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{3.6}$$

The conditional entropy of A conditioned on B in a bipartite state ρ_{AB} is

$$S(A|B)_{\rho_{AB}} = S(AB) - S(B). \tag{3.7}$$

The mutual information in both discrete and continuous cases satisfies a specific inequality

$$I(A;B) \geq 0.$$

This inequality is a consequence of Jensen's inequality. The relative entropy is a very natural way to measure the distance between two probability distributions and as well is a nonnegative quantity $D(\sigma||\rho) \geq 0$. The first result can now be stated explicitly.

Using the property of The von Neumann entropy for product states $\rho_{\mathcal{SR}} = \rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}}$ has the property of additivity. Using this as well as invariance of the entropy under unitary evolution, (**??**), we have

$$S(\rho'_{\mathcal{S}}) - S(\rho_{\mathcal{S}}) + S(\rho'_{\mathcal{R}}) - S(\rho_{\mathcal{R}}) = S(\rho'_{\mathcal{R}}) + S(\rho'_{\mathcal{R}}) - S(\rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}})$$

$$= S(\rho'_{\mathcal{S}}) - S(\rho'_{\mathcal{R}}) - S(\rho'_{\mathcal{SR}}) = I(\mathcal{S}';\mathcal{R}')_{\rho'_{\mathcal{SR}}} \geq 0. \tag{3.8}$$

The last inequality follows from the non-negativity of the mutual information. This proves the following result which will be stated as a theorem.

Theorem. Let $\rho_{\mathcal{SR}} = \rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}} \in \mathcal{B}(\mathbb{C}^{d_S}) \otimes \mathcal{B}(\mathbb{C}^d)$ be a product state on a bipartite finite-dimensional system $\mathcal{SR}$. Let $\mathbf{U} \in \mathcal{B}(\mathbb{C}^{d_S} \otimes \mathbb{C}^d)$ be a unitary operator. Let the state after the evolution be denoted $\rho'_{\mathcal{SR}} = \mathcal{U}(\rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}})\mathcal{U}^\dagger$, with reduced states $\rho'_{\mathcal{S}}$ and $\rho'_{\mathcal{R}}$. Then

$$\big(S(\rho'_{\mathcal{S}}) - S(\rho_{\mathcal{S}})\big) - \big(S(\rho'_{\mathcal{R}}) - S(\rho_{\mathcal{R}})\big) = I(\mathcal{S}';\mathcal{R}') \geq 0. \tag{3.9}$$

Using (**??**) and (**??**), this also takes the form,

$$\Delta = \Delta\, S + I(\mathcal{S}';\mathcal{R}') \geq \Delta\, S. \tag{3.10}$$

Thus the increase of reservoir entropy Δ is greater than the entropy decrease of the system $\Delta\, S \equiv S(\rho_{\mathcal{S}}) - S(\rho'_{\mathcal{S}})$. $\square$

That equality in (**??**) is attained follows from this result. This implies there occurs no total entropy increase, if and only if the final state is in fact a product state. The main objective is now possible, that is, to formulate and prove Landauer's theorem.

Landauer's Principle: Let $\rho_{\mathcal{R}}$ be the thermal state corresponding to Hamiltonian $\mathbf{H} = \mathbf{H}^\dagger$ at inverse temperature $\beta \in [-\infty, \infty]$ and $\rho_{\mathcal{SR}} = \rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}} = \mathcal{B}(\mathbb{C}^{d_{\mathcal{S}}})) \otimes \mathcal{B}(\mathbb{C}^d)$ a product state on a finite-dimensional system $\mathcal{SR}$. Let $\mathbf{U}$ be a unitary operator and let the state after the evolution be $\rho'_{\mathcal{SR}} = \mathbf{U}\big(\rho_{\mathcal{S}} \otimes \rho_{\mathcal{R}}\big)\mathbf{U}^\dagger$ with corresponding reduced states $\rho'_{\mathcal{S}}$ and $\rho'_{\mathcal{R}}$. Then

$$S(\rho_{\mathcal{S}}) - S(\rho'_{\mathcal{S}}) + I(\mathcal{S}';\mathcal{R}') + D(\rho'_{\mathcal{R}};\rho_{\mathcal{R}}) = \beta\,\mathrm{tr}(\mathbf{H}\rho'_{\mathcal{R}}) - \beta\,\mathrm{tr}(\mathbf{H}\rho_{\mathcal{R}}). \tag{3.11}$$

Proof: Suppose first of all that $\beta \in (-\infty, \infty)$. Using results (**??**) and (**??**),

$$\begin{aligned}
\Delta S + I(\mathcal{S}';\mathcal{R}') = S(\rho_{\mathcal{S}}) - S(\rho'_{\mathcal{S}}) + I(\mathcal{S}';\mathcal{R}') = S(\rho'_{\mathcal{R}}) - S(\rho_{\mathcal{R}})\\
= -\mathrm{tr}(\rho'_{\mathcal{R}}\,\log\,\rho'_{\mathcal{R}}) + \mathrm{tr}\Big(\rho_{\mathcal{R}}\log\,\frac{e^{-\beta\mathbf{H}}}{\mathrm{tr}(e^{-\beta\mathbf{H}})}\Big)\\
= -\mathrm{tr}(\rho'_{\mathcal{R}}\log\,\rho'_{\mathcal{R}}) + \mathrm{tr}(\rho_{\mathcal{R}}(-\beta\,\mathbf{H} - \mathbf{1}\,\log\mathrm{tr}(e^{-\beta\mathbf{H}})))\\
= -\mathrm{tr}(\rho'_{\mathcal{R}}\,\log\rho'_{\mathcal{R}}) - \beta\,\mathrm{tr}(\mathbf{H}\rho_{\mathcal{R}}) - \log\,\mathrm{tr}(e^{-\beta\mathbf{H}})
\end{aligned}$$

Add and subtract $\beta\,\mathrm{tr}\,(\mathbf{H}\rho'_{\mathcal{R}})$ on the right side of this result, it follows that

$$\Delta S + I(\mathcal{S}';\mathcal{R}') = \mathrm{tr}\,(\rho'_{\mathcal{R}}\log\rho'_{\mathcal{R}}) - \beta\,\mathrm{tr}\,(\mathbf{H}\rho_{\mathcal{R}}) - \log\,\mathrm{tr}(e^{-\beta\mathbf{H}}) + \beta\,\mathrm{tr}\,(\mathbf{H}\rho'_{\mathcal{R}}) - \beta\,\mathrm{tr}(\mathbf{H}\rho'_{\mathcal{R}})$$

$$\begin{aligned}
= \beta\mathrm{tr}\,(\mathbf{H}(\rho'_{\mathcal{R}} - \rho_{\mathcal{R}})) - \mathrm{tr}(\rho'_{\mathcal{R}}\log\,\rho'_{\mathcal{R}}) - \log\mathrm{tr}(e^{-\beta\mathbf{H}}) - \mathrm{tr}\,(\beta\mathbf{H}\rho'_{\mathcal{R}})\\
= \beta\,\mathrm{tr}(\mathbf{H}(\rho'_{\mathcal{R}} - \rho_{\mathcal{R}})) - \mathrm{tr}(\rho'_{\mathcal{R}}\log\rho'_{\mathcal{R}}) + \mathrm{tr}\Big(\rho'_{\mathcal{R}}\log\,\frac{e^{-\beta\mathbf{H}}}{\mathrm{tr}\,\big(e^{-\beta\mathbf{H}}\big)}\Big)\\
= \beta\,\Delta Q - D\Big(\rho'_{\mathcal{R}}||\log\,\frac{e^{-\beta\mathbf{H}}}{\mathrm{tr}\,\big(e^{-\beta\mathbf{H}}\big)}\Big) = \beta\Delta Q - D(\rho'_{\mathcal{R}}||\rho_{\mathcal{R}}).
\end{aligned}$$

Solving this equation for $\beta\,\Delta Q$, equation (**??**) appears

$$\beta\,\Delta\,Q = \Delta\,S + I(\mathcal{S}';\mathcal{R}') + D(\rho'_{\mathcal{S}}||\rho_{\mathcal{R}}). \tag{3.12}$$

Consider the special case $\beta = +\infty$, then

$$\rho_{\mathcal{R}} = \frac{\pi_g}{\dim(\pi_g)}. \tag{3.13}$$

Equation (**??**) is the normalized projector onto the ground state of $\mathbf{H}$. This implies $\mathrm{tr}\,(\mathbf{H}\,\rho'_{\mathcal{R}}) \geq \mathrm{tr}(\mathbf{H}\rho_{\mathcal{R}})$ thus $\Delta\,Q \geq 0$. If $\Delta\,Q = 0$, then $\rho'_{\mathcal{R}}$ is supported in the ground state space as well. Thus one may conclude

$$S(\rho'_{\mathcal{R}}) - S(\rho_{\mathcal{R}}) = -\mathrm{tr}\,(\rho'_{\mathcal{R}} \log\, \rho'_{\mathcal{R}}) - \log \dim(\pi_g) = -D(\rho'_{\mathcal{R}}||\rho_{\mathcal{R}}).$$

This implies that both sides of (**??**) vanish. If $\Delta Q > 0$, $\rho'_{\mathcal{R}}$ has nonzero support beyond that of ground state space of $\mathbf{H}$. Outside the support of $\rho_{\mathcal{R}}$, $D(\rho'_{\mathcal{R}}||\rho_{\mathcal{R}}) = +\infty$, and both sides of (**??**) are equal to each other again. The result for $\beta = -\infty$ follows from this process by using the transformation $\mathbf{H} \rightarrow -\mathbf{H}$ and $\beta \rightarrow -\beta$. The fact that $\beta\,\Delta\,Q \geq S$ comes about from the fact that mutual information and the relative entropy are both non-negative quantities. □

The issue of how to handle systems and reservoirs which do not have finite-dimensional Hilbert spaces as described so far may be addressed. Some of the quantities that appear in Landauer's principle may not be defined or may require an alternate definition. If the initial and final system entropies turn out to be infinite since $\Delta\,S$ is ill-defined. However, if the Hilbert space is separable a portion of this can be carried over. To do this, assume the initial state of S is a normal state $\rho_{\mathcal{S}}$ such that $S(\rho_{\mathcal{S}}) < \infty$, and a semi-bounded Hamiltonian $\mathbf{H}$ is given for $\mathcal{R}$ so at inverse temperature $\beta \in (0, \infty]$, the thermal state $\rho_\beta = \rho_{\mathcal{R}}$ exists and has finite energy. These two conditions are expressed in the mathematically equivalent way $\mathrm{tr}(e^{-\beta \mathbf{M}}) < \infty$ and $\mathrm{tr}\,(\mathbf{H}e^{-\beta\mathcal{H}}) < \infty$. They imply entropy $S(\rho_{\mathcal{R}})$ is also finite. For a unitary $\mathbf{U}$ the first result remains valid as well so all quantities are finite, except the case $I(\mathcal{S}'; \mathcal{R}') = +\infty$ may arise.

Studying the derivation of Landauer's principle we have $\beta\,\Delta Q \geq -\infty$. This must be as $\mathbf{H}$ is semi-bounded and $\rho_{\mathcal{R}}$ has finite energy. Moreover, $\Delta\,S = -\infty$ implies that $\beta\Delta\,Q = \infty$, which is due to $D(\rho'_{\mathcal{R}}||\rho_{\mathcal{R}}) \geq 0$. The form of Landauer's principle (**??**) under equality holds in this instance when rules of calculus are used in addition for ∞. In the possibly ambiguous case $\Delta\,S = -\infty$, $\beta\,\Delta\,Q \geq \Delta\,S$ always holds. Even if a process like such as the one discussed above with infinite $S(\rho_{\mathcal{S}})$ and finite $S(\rho'_{\mathcal{S}})$, an infinite $\Delta\,S = \infty$ of entropy is erased from $\mathcal{S}$. It follows $S(\rho'_{\mathcal{R}}) = \Delta = \beta\Delta\,Q = \infty$. In this case, Landauer's bound holds as well.

4 Physical Model Examples

Some realistic physical examples are presented to clarify and give a more physical picture associated with the rigorous approach which has just been presented. The first example follows.

(i) Let both the system $\mathcal{S}$ and reservoir $\mathcal{R}$ have bipartite structure and match in the dimension of one subsystem: $\mathbb{C}^{d_{\mathcal{S}}} = \mathbb{C}^{d_{sw}} \otimes \mathbb{C}^{d_{\mathcal{S}_2}}$, $\mathbb{C}^{d} = \mathbb{C}^{d_{sw}} \otimes \mathbb{C}^{d_{\mathcal{R}_2}}$, where $d_{sw}, d_{\mathcal{S}_2}$, $d_{\mathcal{R}_2} \in \mathbb{N}$. The reduced states of $\rho_{\mathcal{S}}$, respectively $\rho_{\mathcal{R}}$, are denoted with respect to these bipartitions as $\rho_{\mathcal{S}_1}$ and $\rho_{\mathcal{R}_2}$. The initial state for the system ρ can be chosen arbitrarily. Also for our purposes, $\rho_{\mathcal{R}}$ may be chosen in a virtually arbitrary manner. This is because every full rank state $\rho_{\mathcal{R}}$ can be written as a thermal state pertaining to Hamiltonian $\mathbf{H} = -\log \rho_{\mathcal{R}}$ at inverse temperature $\beta = 1$. Let the unitary matrix for enacting the process be given by $\mathbf{U} = \mathbb{F}_{\mathcal{S}_1,\mathcal{R}_1} \otimes \mathbb{I}_{\mathcal{S}_2,\mathcal{R}_2}$. This is composed of the unitary flip $\mathbb{F}$ which swaps the two subsystems $\mathcal{S}_1$ and $\mathcal{R}_1$ of dimension d_{sw} acts according to

$$\mathbb{F}_{\mathcal{S}_1,\mathcal{R}_1}\big(|\psi_1\rangle \otimes |\psi_2\rangle\big) = |\psi_2\rangle \otimes |\psi_1\rangle. \tag{4.1}$$

This allows $\mathcal{S}_2, \mathcal{R}_2$ to remain unaltered. The following quantities can be determined explicitly for this kind of evolution

$$\Delta\, S = S(\rho_{\mathcal{S}}) - S(\rho_{\mathcal{S}_2}) - S(\rho_{\mathcal{R}_1}), \qquad \Delta = -S(\rho_{\mathcal{R}}) + S(\rho_{\mathcal{R}_2}) + S(\rho_{\mathcal{S}_1}),$$

$$\Delta\, Q = \mathrm{tr}[\mathbf{H}(\rho_{\mathcal{S}_1} \otimes \rho_{\mathcal{R}_2} - \rho_{\mathcal{R}})], \qquad I(\mathcal{S}'; \mathcal{R}') = I(\mathcal{S}_1; \mathcal{S}_2) + I(\mathcal{R}_1; \mathcal{R}_2). \tag{4.2}$$

(ii) Consider the following second model. Suppose there is a two-state system with a 2×2 density matrix with $\rho_{\mathcal{S},i} = 1/2$ on the main diagonal. The heat reservoir is in thermal equilibrium initially with a state with density matrix (**??**). Suppose $\mathcal{R}$ has just a finite number of states and imagine it to be in a definite energy eigenstate $|E_n\rangle$ and eigenvalue E_n so

$$P_n = \frac{e^{-\beta E_n}}{\sum_i e^{-\beta\, E_n}} = \frac{1}{Z}\, e^{-\beta E_n}. \tag{4.3}$$

When the erasure is successful, the system ends up in the pure state corresponding to one. The reservoir will be in some statistical state. This may typically be described by a density matrix which is not diagonal in the energy eigenbasis. To clarify the calculation, suppose the reservoir is weakly coupled to an environment causing off-diagonal elements to vanish such as due to decoherence and not changing the diagonal. There is a classical element associated to this. Thus the two-state system starts in zero or one but ends in the one, while the reservoir starts in some state $|n\rangle$ and ends up in $|m\rangle$. For $i, f = 0, 1$, let $\rho_{\mathcal{S},i}$ and $\rho_{\mathcal{S},f}$ correspond to the probability distribution of those states. Define the functional F as

$$F = \log(\rho_{\mathcal{S},i}) - \log(\rho_{\mathcal{S},f}) - \beta(E_n - E_m). \tag{4.4}$$

Using F from (**??**), calculate $\langle e^{-F}\rangle$. This is subjected to a unitary transformation $\mathbf{U}$ which has matrix elements $U_{f,m;i,n}$

$$\langle e^{-F}\rangle = \sum_{n,m,i,f} \rho_{S,i} P_n \, |U_{f,m;i,n}|^2 \, e^{-\log \rho_{\mathcal{S},i} + \log \rho'_{\mathcal{S},f}) + \beta(E_n - E_m)}$$

$$= \sum_{n,m,i,f} \rho'_{\mathcal{S},f} |U_{f,m;i,n}|^2 \frac{e^{-\beta E_n}}{Z} \cdot e^{-\beta\, E_m} \cdot e^{\beta E_m} = \frac{1}{Z} \sum_{f,m} \rho'_{\mathcal{S},f} \, e^{-\beta E_m} \sum_{i,n} |U_{f,m,i,n}|^2. \tag{4.5}$$

Equation (**??**) has been obtained by substituting (**??**). As $\mathbf{U}$ is a unitary matrix, the sum of the squares of the absolute values of its elements in a column or row is one. Thus $|U_{f,m,i,n}|^2$ is the probability to find the system and reservoir in $|f\rangle$ and $|m\rangle$ respectively given initial states $|i\rangle$ and $|n\rangle$. This implies that

$$\langle e^{-F}\rangle = 1. \tag{4.6}$$

Now concavity of the exponential function yields the inequality

$$-\langle F\rangle \leq 0.$$

Substituting functional (**??**) into this result, we obtain

$$\langle \log(\rho_{\mathcal{S},i})\rangle + \langle \log(\rho'_{\mathcal{S},f}\rangle + \langle \beta(E_n - E_m)\rangle \leq 0. \tag{4.7}$$

Substituting for these remaining quantities, it is deduced that under the identification $Q = E_m - E_n$ this Q is the heat dissipated into the heat reservoir

$$-\log(\frac{1}{2}) - \beta \, \langle Q\rangle \leq 0.$$

Equivalently, this can be put in the form,

$$\beta \langle Q \rangle \geq \log(2). \tag{4.8}$$

Remarkably, studying both the system and heat reservoir, work can be defined as in the classical case, that is, $W = \Delta E_{heat} + \Delta E_{system}$. This equation is valid under the assumption that the interaction energy between the heat reservoir and two-state system is very small, just as in the continuous classical case.

5 References

[1] Maxwell's Demon 2. Entropy, Classical and Quantum Information, Computing, Eds. H. S. Leff and A. F. Rex IOP Publishing Ltd., Bristol, 1995.

[2] L. Szilard, Z. f. Physik, **53**, 840-856, (1929).

[3] P. Bracken, Central European Journal of Physics, **12**, 1-8, (2014).

[4] L. van Hove, Physica **A 23**, 144-480, (1957).

[5] D. Reeb and H. Wolf, New Journal of Physics, bf 16, 103011, (2014).

[6] C. Jarzynski, J. Stat. Phys. **96**, 415-427, (1999).

[7] C. Jarzynski, Annu. Rev. Condensed Matter Physics, **2**, 329-351, (2011).

[8] B. Piechocinska, Phys.Rev. **A 61**, 062314, (2000).

[9] J. Anders and V. Giovannetti, New J. Physics, **15**, 033022, (2013).

[10] C. Bennett, Int. J. Theor. Physics, **21**, 905-940, (1982).

[11] A. Wehrl, Rev. Mod. Physics, **50**, 221-260, (1978).