情報, 熱力学, (そして統計力学)

丸山耕司
(大阪市立大学)

KM, F. Nori, and V. Vedral, Rev. Mod. Phys. 81, 1 (2009)
KM, 数理科学 2012年3月号

Outline

Introduction

- Shannon entropy
- Maxwell's demon paradox

Landauer's erasure principle

Quantum case

Some implications of the second law

Demon and data compression

Summary

Shannon entropy

How can information be quantified?

- The more valuable, the more information.

  useful

  surprising

  smaller probability

  Information $I(j)$ in message j :
  decreasing function of $p_j = prob(j)$

Roughly, Information = the degree of (our) ignorance

- Information of a joint event = Sum of information of each event

$$I(j,k) = I(j) + I(k)$$

$$(p(j,k) = p_j p_k)$$

Shannon entropy

To satisfy these naive requirements,

- $I(j)$ is a decreasing function of
- $I(j,k) = I(j) + I(k) \quad (p(j,k) = p_j p_k)$

lets quantify information as $I(p_j) = -\log_2 p_j$ [bit]

The average information per single alphabet for a given probability distribution $\{p_i\}$, where $\sum_i p_i = 1$, is

$$H(p) := -\sum_i p_i \log_2 p_i. \qquad \text{(Shannon) entropy}$$

A sequence of N binary numbers, whose information content is $H(p)$, can be re-expressed with a $NH(p)$ bit sequence.

Consider a sequence of N bit string, where 0 and 1 appear with probabilities of 1-1/N and 1/N, respectively.

$$\overbrace{0 \ldots 0\, 0 \ldots\ldots 0\, 1}^{N}$$

$$0 \ldots 0\, 0 \ldots\ldots 1\, 0$$

$$0 \ldots 0\, 1 \ldots\ldots 0\, 0$$

$$1 \ldots 0\, 0 \ldots\ldots 0\, 0$$

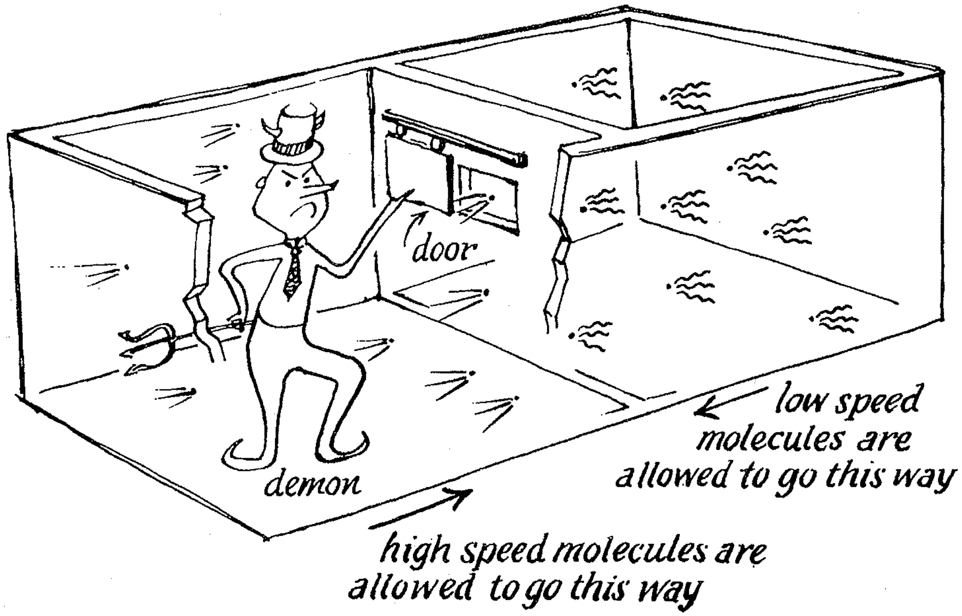
One of these sequences would occur almost for sure (with high probability).

$\log_2 N$ bits would be sufficient to specify the sequence that occured.

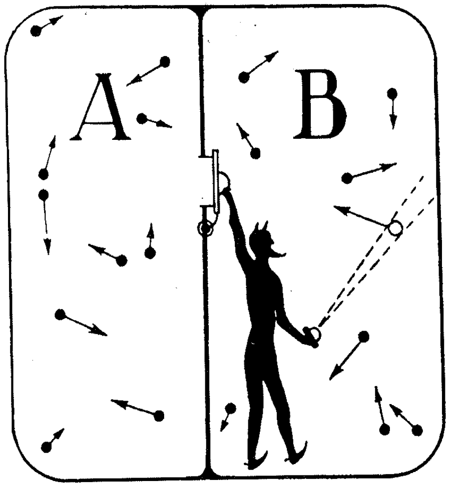
ex. if $N = 16$, 0000, 0001, ..., 1111.

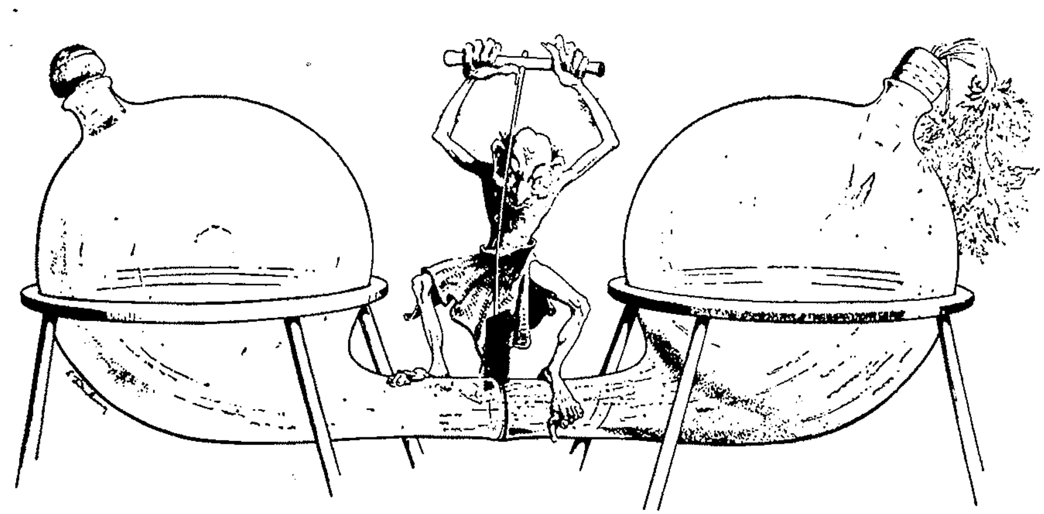
In fact, $\log_2 N$ is a good approximation for $NH(1/N)$.

Maxwell's demon

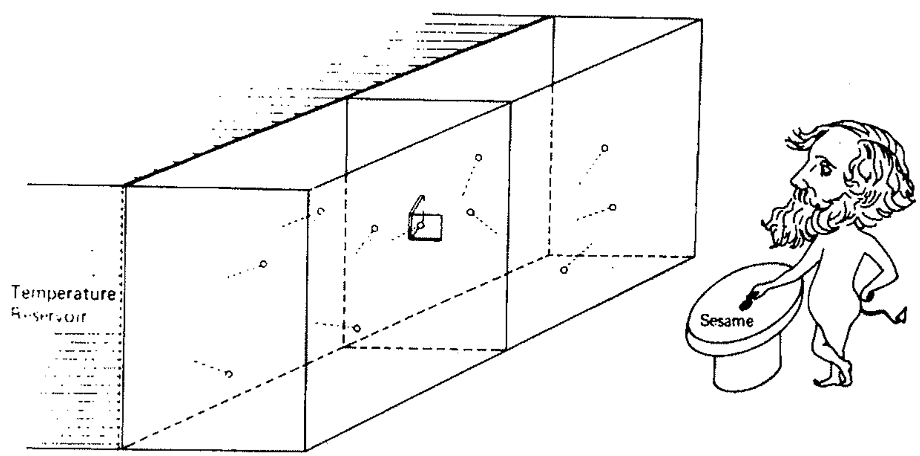


Maxwell's demon





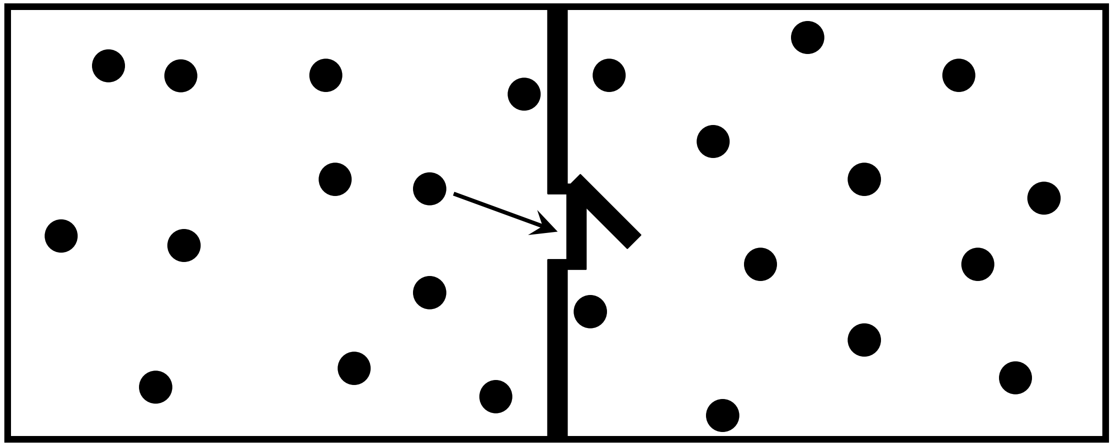
Maxwell's demon at work



A Maxwell demon controlling a door between two chambers each initially at temperature T_1 and pressure P_1

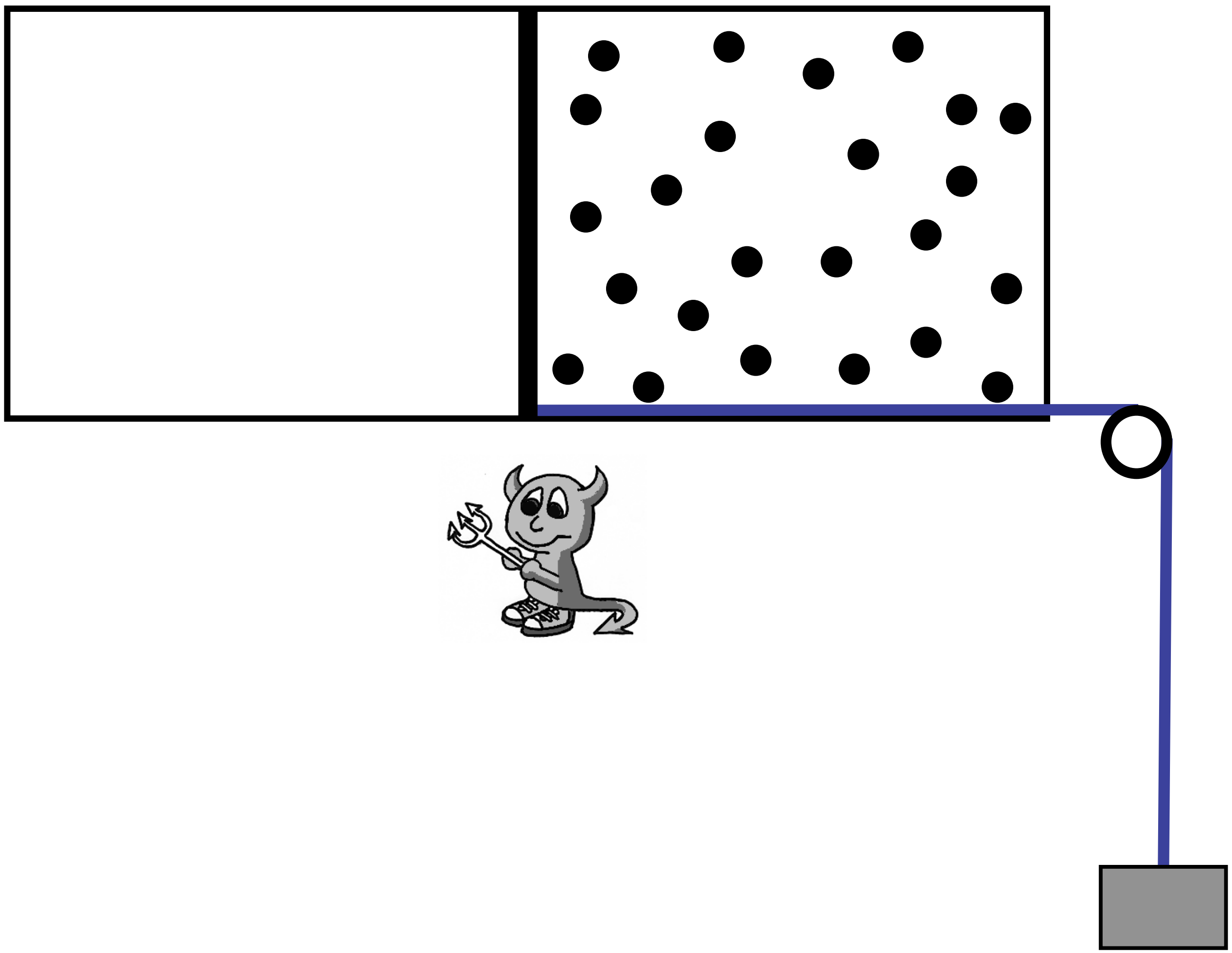
Maxwell's demon

Initial state

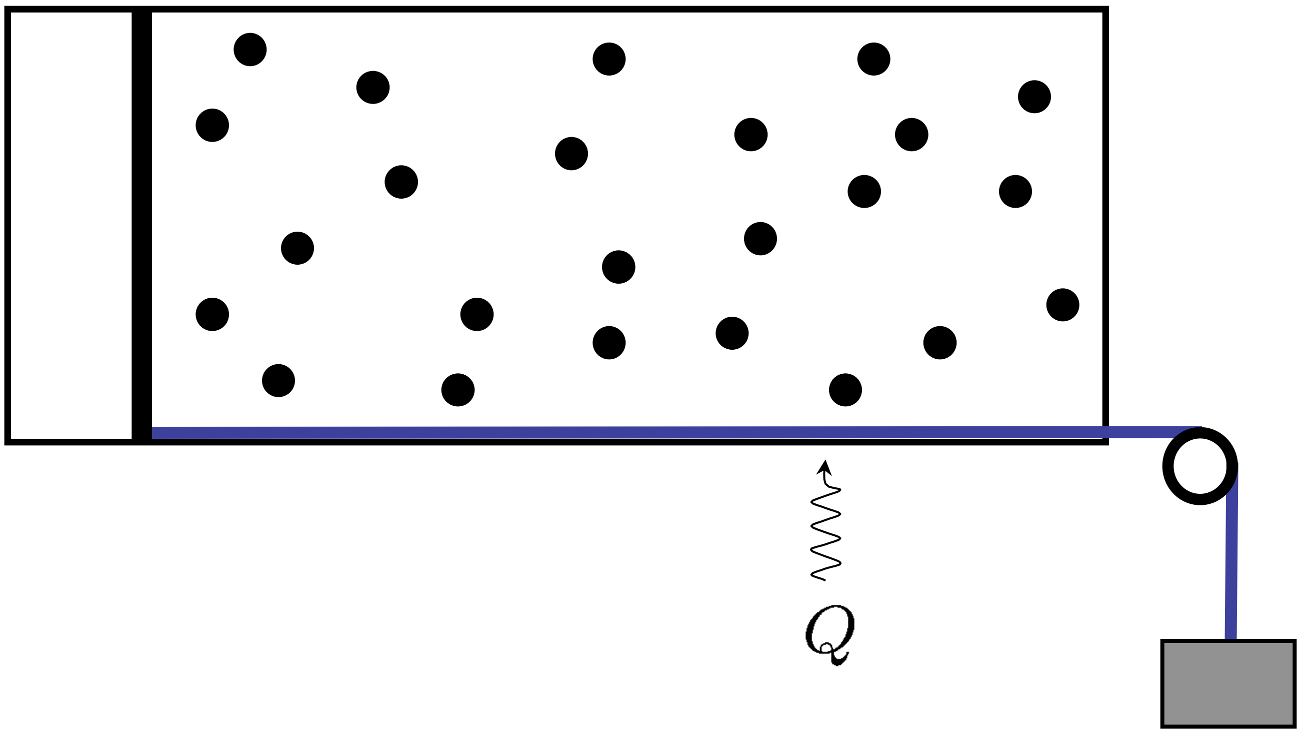


K. Maruyama

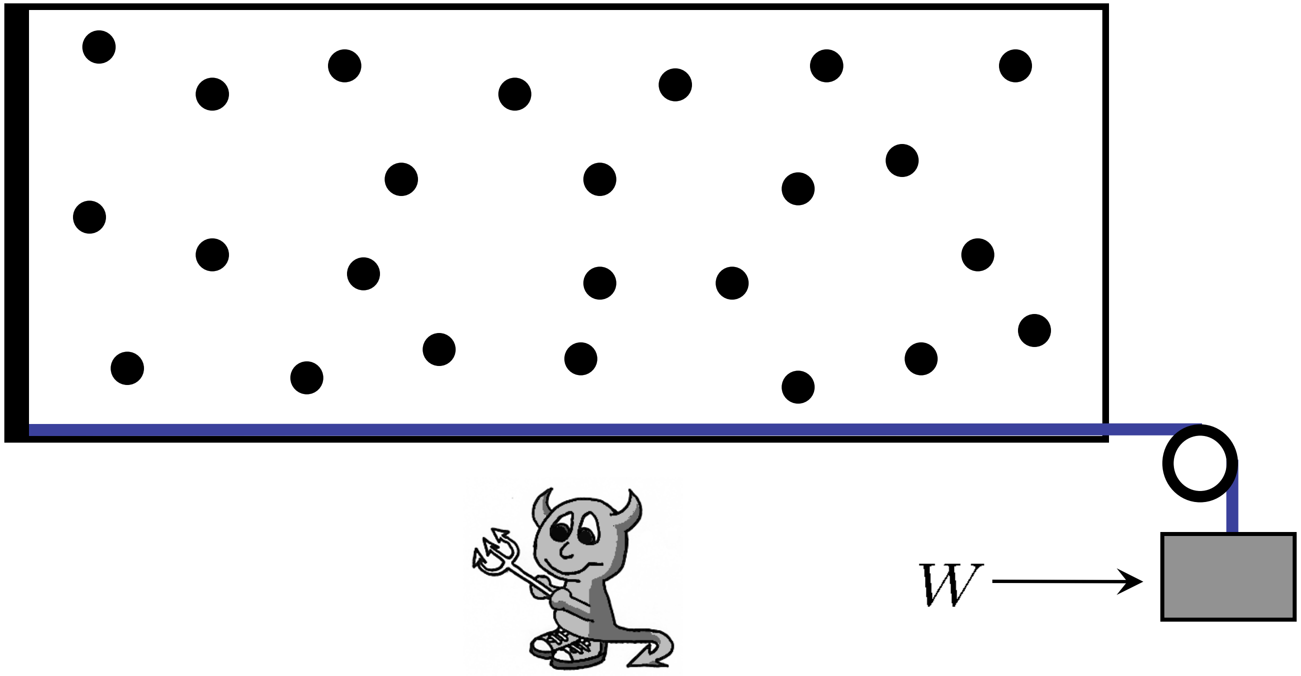
Maxwell's demon



Isothermal expansion

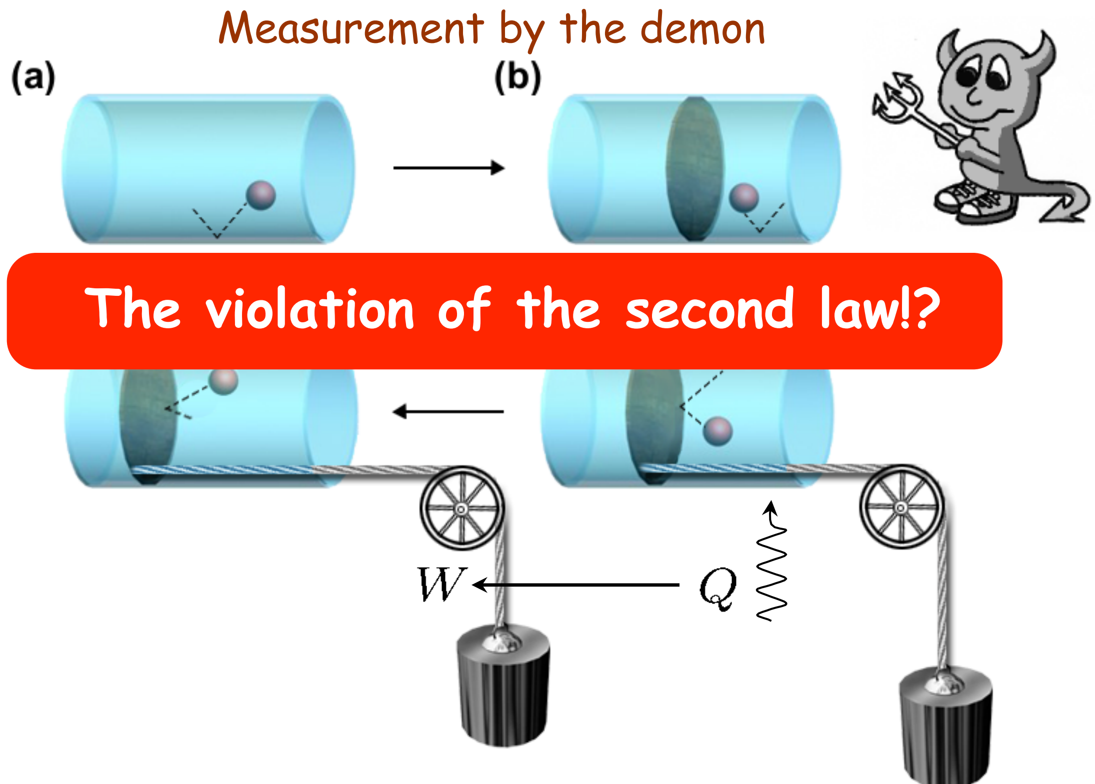


Initial state again!



K. Maruyama

Maxwell's demon paradox a la Szilard



The violation of the second law!?

Isothermal expansion at temperature T

Perfect conversion of heat Q into mechanical work $W = kT \int_{V/2}^{V} \frac{1}{V} dV = kT \ln 2.$

Outline

Introduction

- Shannon entropy
- Maxwell's demon paradox

Landauer's erasure principle

Quantum case

Some implications of the second law

Demon and data compression

Summary

Where's the dissipation (entropy increase)?

Szilard (1929): "Must be in the measurement process."

Brillouin (1951): "Yes, indeed. It is in the measurement process."
after calculating the entropy change caused by laser-based measurement.

Brillouin : Put thermodynamic and information entropies in the same equation.

Second law of thermodynamics applicable to the 'sum' of entropies.

"Negentropy + information never increases."

Conjectured the equivalence between the two entropies.

Landauer's principle

Landauer (1961): "Information is physical."

One-to-one correspondence

Logical "0" ⟺ Physical "0" state $\vec{x}_0$

Logical "1" ⟺ Physical "1" state $\vec{x}_1$

$$\vec{x}_i = (x_i^0, x_i^1, ..., x_i^N)$$

A set of parameters defining the physical state "i".

Logically irreversible process (many-to-one mapping)

⇩

Physically irreversible process (reduction of degrees of freedom)

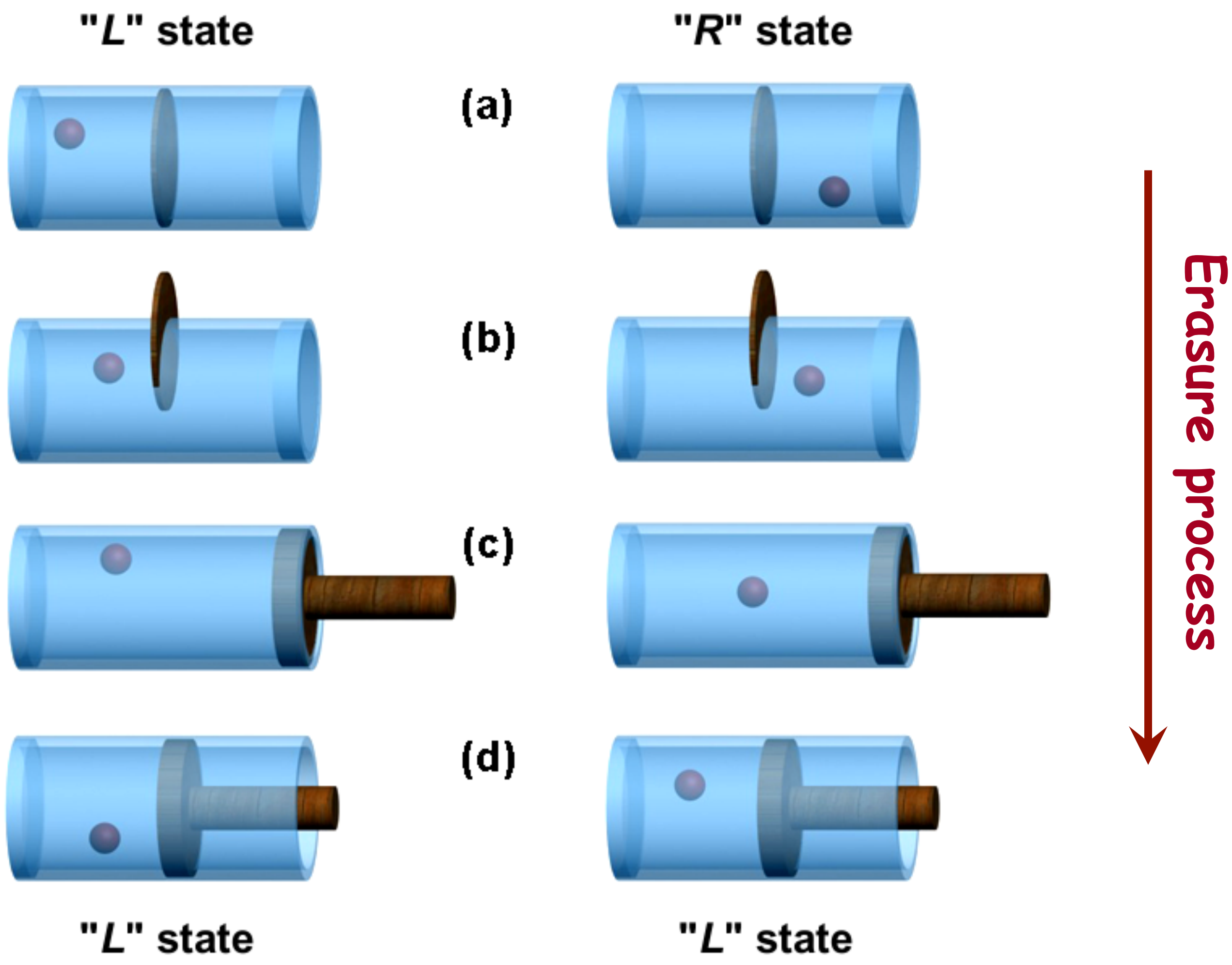
⇩

Dissipation of energy into the environment

Erasing information entails entropy increase in the environment.

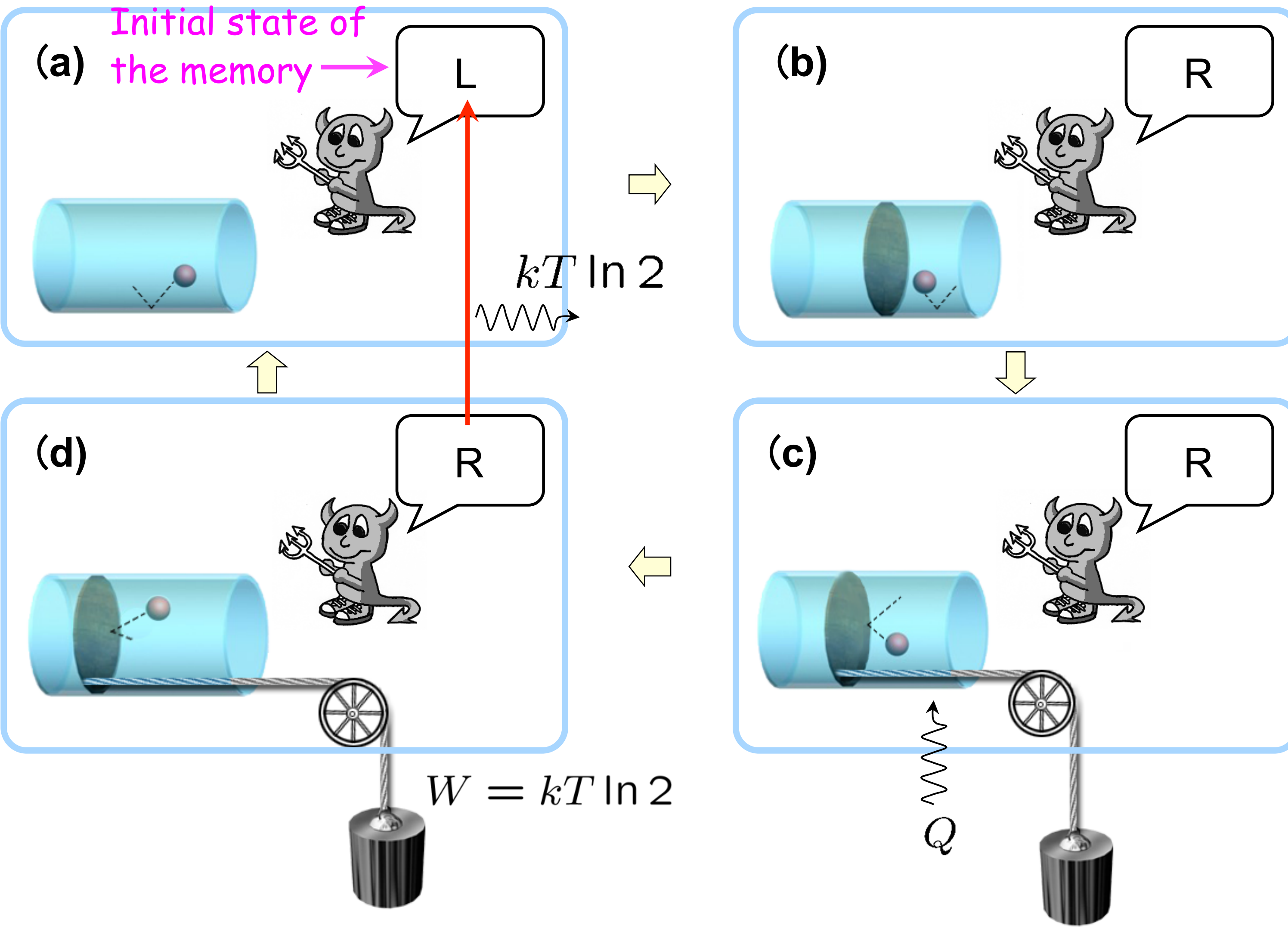
(Landauer's erasure principle)

Modelling a memory by a one-molecule gas



Landauer's principle

The demon's memory (in Szilard's engine) is also a physical object.

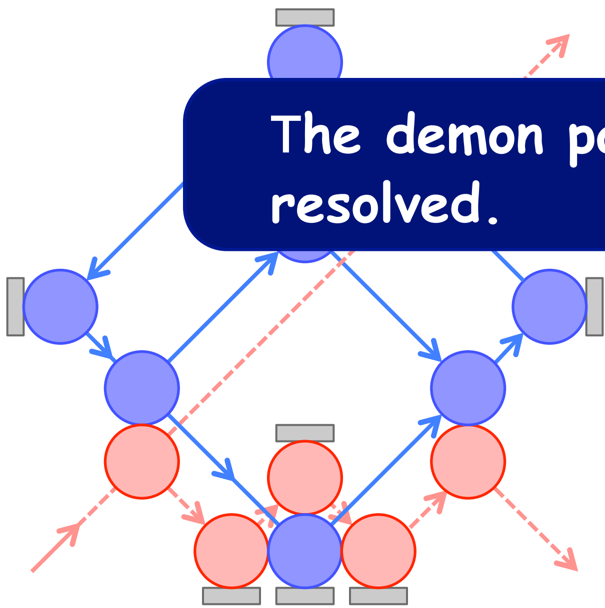


K. Maruyama

Measurements can be performed reversibly. (Bennett, 1982)

examples:

measurement (detection) of the blue ball without disturbing its motion



reversible copying of information using a small magnet (Bennett, 1982)

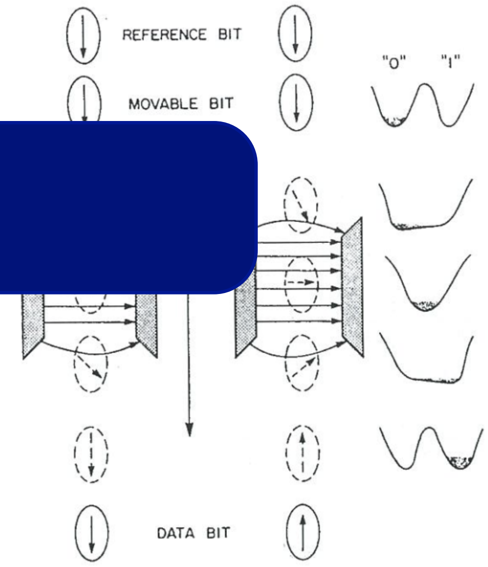
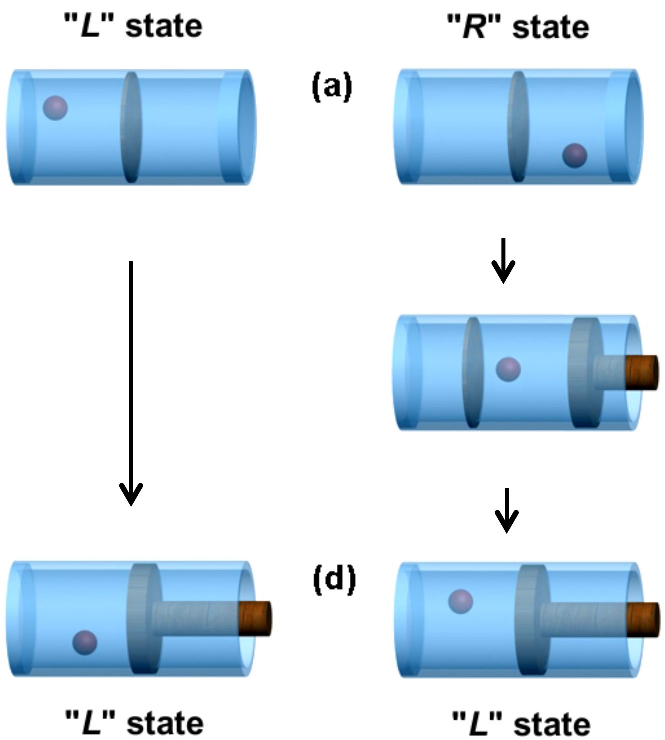


Fig. 14. Reversible copying using a one-domain ferromagnet. The movable bit, initially zero, is mapped into the same state as the data bit (zero in left column; one in center column). Right column shows how the probability density of the movable bit's magnetization, initially concentrated in the "0" minimum, is deformed continuously until it occupies the "1" minimum, in agreement with a "1" data bit.

The demon paradox now resolved.

criticisms

- No work needed for erasure!



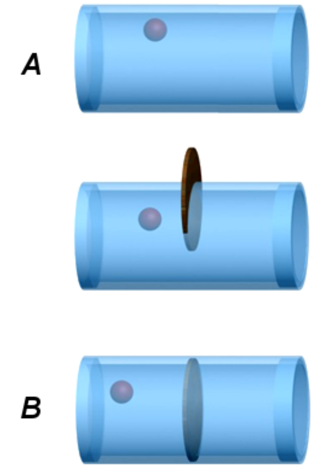
- This operation leaves the information on the initial state

- Violation of the 2nd law by wall-insertion

$$\int_A^B \frac{d'Q}{T} \leq S(B) - S(A)$$

$$d'Q = 0$$

$$S(B) - S(A) = -k \ln 2$$



lower entropy ⇨ larger free energy

- N molecules in the cylinder
  ⇨ no entropy change

- Need *information* to extract work
  ⇨ demon comes in and erasure necessary

an experimental verification of Landauer

A. Berut et al., Nature 483, 187 (2012).

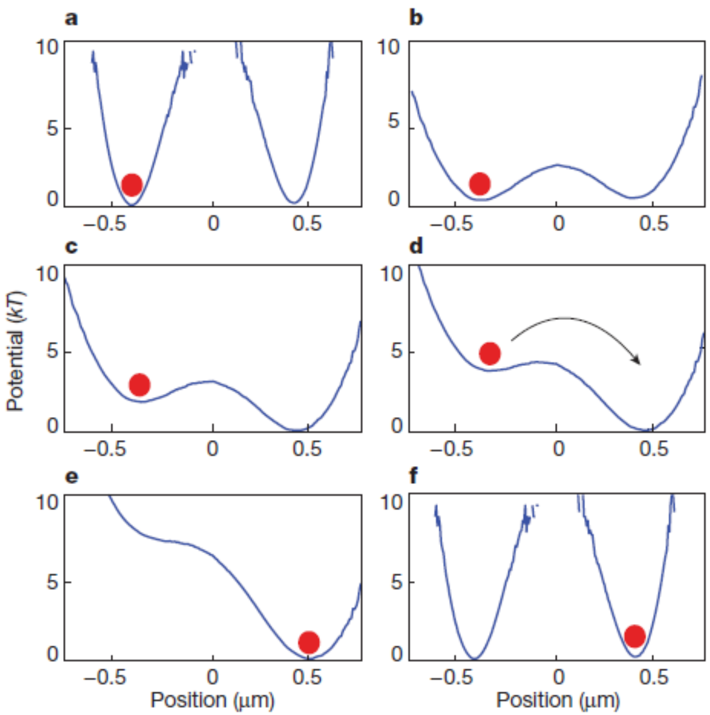
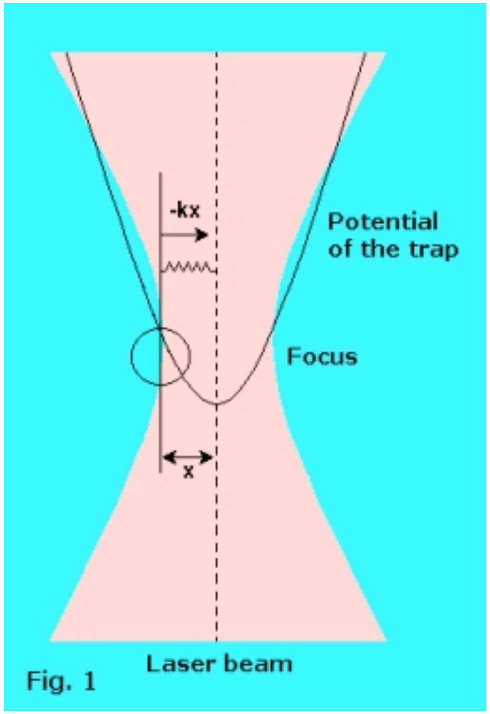


Figure 1 | The erasure protocol used in the experiment. One bit of information stored in a bistable potential is erased by first lowering the central barrier and then applying a tilting force. In the figures, we represent the transition from the initial state, 0 (left-hand well), to the final state, 1 (right-hand well). We do not show the obvious 1 → 1 transition. Indeed the procedure is such that irrespective of the initial state, the final state of the particle is always 1. The potential curves shown are those measured in our experiment (Methods).



silica-bead-trap by optical tweezer

an experimental verification of Landauer

A. Berut et al., Nature 483, 187 (2012).

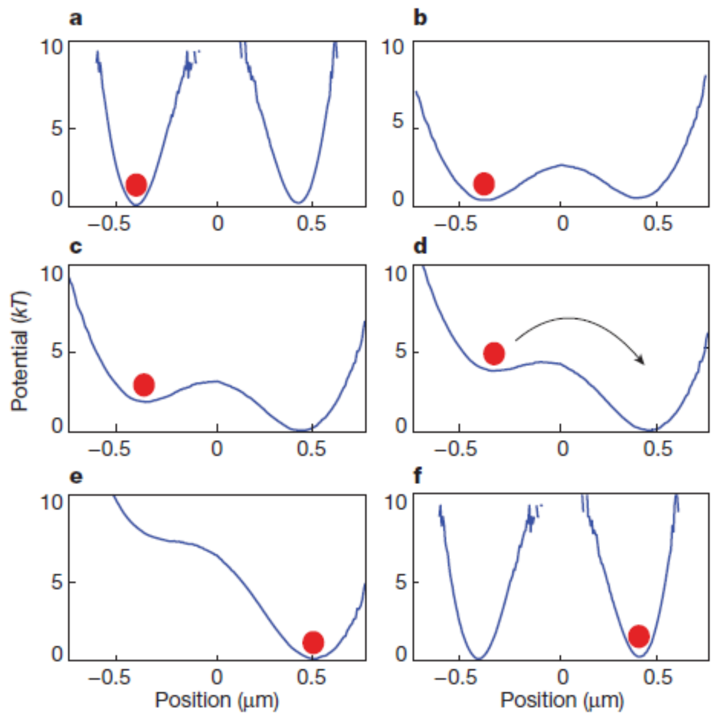


Figure 1 | The erasure protocol used in the experiment. One bit of information stored in a bistable potential is erased by first lowering the central barrier and then applying a tilting force. In the figures, we represent the transition from the initial state, 0 (left-hand well), to the final state, 1 (right-hand well). We do not show the obvious 1 → 1 transition. Indeed the procedure is such that irrespective of the initial state, the final state of the particle is always 1. The potential curves shown are those measured in our experiment (Methods).

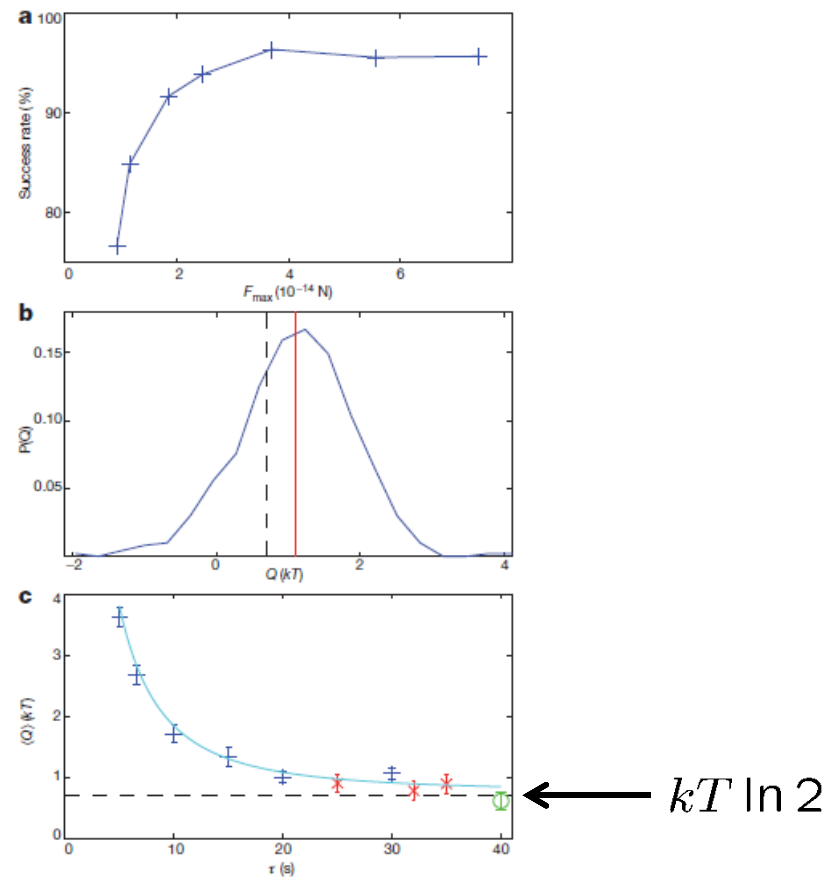


Figure 3 | Erasure rate and approach to the Landauer limit. **a**, Success rate of the erasure cycle as a function of the maximum tilt amplitude, F_{max}, for constant $F_{max}\tau$. **b**, Heat distribution P(Q) for transition 0 → 1 with $\tau = 25$ s and $F_{max} = 1.89 \times 10^{-14}$ N. The solid vertical line indicates the mean dissipated heat, $\langle Q \rangle$, and the dashed vertical line marks the Landauer limit, $\langle Q \rangle_{Landauer}$. **c**, Mean dissipated heat for an erasure cycle as a function of protocol duration, τ, measured for three different success rates, r: plus signs, $r \geq 0.90$; crosses, $r \geq 0.85$; circles, $r \geq 0.75$. The horizontal dashed line is the Landauer limit. The continuous line is the fit with the function $[A\exp(-t/\tau_K) + 1]B/\tau$, where τ_K is the Kramers time for the low barrier (Methods). Error bars, 1 s.d.

Outline

Introduction

Shannon entropy

Maxwell's demon paradox

Landauer's erasure principle

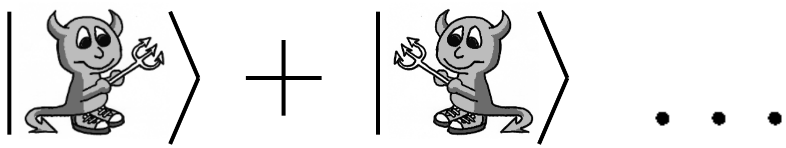
Quantum case

Some implications of the second law

Demon and data compression

Summary

What if we go into quantum regime?



Classical information encoded in quantum system

$|I\rangle$ $|l\rangle$ $|o\rangle$ $|v\rangle$ $|e\rangle$ $|y\rangle$ $|o\rangle$ $|u\rangle$ $|\heartsuit\rangle$

erasure of classical info in quantum states

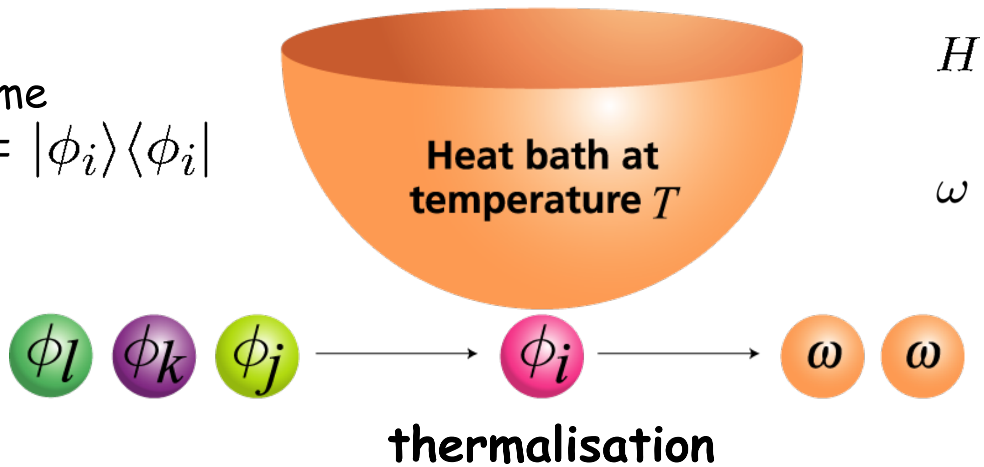
An information source generates $i \in \{1, 2, ..., n\}$ with probability p_i .

Encode i in quantum states ρ_i .

How much entropy increase to erase information in $\{p_i, \rho_i\}$?

Lets thermalise the memory system so that all info will be lost.

Assume
$\rho_i = |\phi_i\rangle\langle\phi_i|$



H : Hamiltonian for particles

$$\omega = \frac{e^{-\beta H}}{Z} = \sum_j q_j |e_j\rangle\langle e_j|$$

(Lubkin, 1987; Vedral, 2000)

von Neumann entropy

$$S(\rho) := -\mathrm{Tr}[\rho \log_2 \rho]$$

To make a long story short,

$$\Delta S_{\mathrm{erasure}} \geq k \ln 2 S(\rho),$$

where $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|.$

Or

$$\Delta S_{\mathrm{erasure}} \geq k \ln 2[S(\rho) - \sum_i p_i S(\rho_i)],$$

if $\rho = \sum_i p_i \rho_i$, where ρ_i are mixed states.

(M. B. Plenio, PLA 1999; KM et al., JPA 2005)

Outline

Introduction

- Shannon entropy
- Maxwell's demon paradox

Landauer's erasure principle

Quantum case

Some implications of the second law

Demon and data compression

Summary

the 2nd law is powerful!

The second law can be lead to a number of interesting implications in other areas of science that look unrelated at first sight.

- The distinguishability of quantum states
- The linearity of quantum mechanics
- Superposition principle
- Einstein equation
- Quantum channel capacity (Holevo bound)

some implications of the 2nd law

- The linearity of the time evoulution of quantum states (Peres, PRL 1989)

$$\rho = p|\phi\rangle\langle\phi| + (1-p)|\psi\rangle\langle\psi|$$

$f := |\langle\phi|\psi\rangle|^2$ (fidelity) Eigenvalues: $\lambda_\pm = \frac{1}{2} \pm \left(\frac{1}{4} - p(1-p)(1-f)\right)^{\frac{1}{2}}$

von Neumann entropy : $S(\rho) = -\lambda_+ \log \ldots \lambda_-$

⇨ $dS/df \le 0$ for ... ⇨ $df/dt \le 0$ in order for entropy to increase

The 2nd law ⇨ Unitarity

Let $\{|\phi_k\rangle\}$ be a complete orthogonal set ⇨ $\sum_k |\langle\phi_k|\psi\rangle|^2 = 1$

⇨ If $\exists m,\ f_m(t) = |\langle\phi_m(t)|\psi(t)\rangle|^2 < f_m(0) \Rightarrow \exists n,\ f_n(t) > f_n(0)$

⇨ $f = |\langle\phi|\psi\rangle|^2$ is constant for any $|\phi\rangle$ and $|\psi\rangle$

⇨ Time evolution needs to be either unitary or ~~anti-unitary~~.

some implications of the 2nd law

- Einstein equation (Bekenstein, PRD 1973; Jacobson, PRL 1995)

Black holes : exact solutions of the Einstein equation

⇨ no randomness involved ⇨ **zero entropy**

⇨ What if we pour hot (high-entropy) coffee into a black hole?

Still zero? **The violation of the second law???**

Bekenstein : "The black hole entropy $\propto$ the area of the event horizon."

Hawking : $S_{BH} = \dfrac{kA}{4l_P^2}$ $l_P = \sqrt{\dfrac{G\hbar}{c^3}}$: Planck length

some implications of the 2nd law

- Einstein equation (Bekenstein, PRD 1973; Jacobson, PRL 1995)

Jacobson : in thermodynamics

$\delta Q = TdS$ ⟹ Equation of state

⇧

$S = S(E, V)$ $f(g_{\mu\nu}, \partial_\alpha g_{\mu\nu}, \partial_\alpha \partial_\beta g_{\mu\nu}) = T_{\mu\nu}$

The Einstein equation ⇦ The equation of state for gravitational field

δQ ⟶ Energy flow across the horizon

T ⟶ The Unruh temperature

dS ⟶ Surface area of the horizon $dS = \eta dA$

$R_{\mu\nu}$ appears in an equation for the volume change in a Riemannian manifold.

⇨

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{2\pi k}{\hbar \eta} T_{\mu\nu}$$

$$\eta = \frac{k}{4\hbar G} = \frac{k}{4 l_P^2}$$

Einstein's field equation !

Outline

Introduction

- Shannon entropy
- Maxwell's demon paradox

Landauer's erasure principle

Quantum case

Some implications of the second law

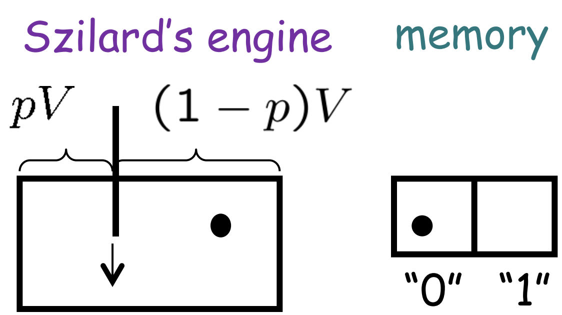
Demon and data compression

Summary

Landauer-Bennett's principle strongly suggests the equivalence between the two entropies.

information theoretic and thermodynamic ones

Show the equivalence, using a **thermodynamic** operational model (of erasure) for **any** probability distribution.



$$W_{\text{sze}} = kT \ln 2 H(p) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$$

nonequiprobability distribution

What is the optimal work consumption to erase the information of $H(p)$?

We'll use (Shannon's) data compression. (as you might have guessed)

N cycles of the Szilard's engine

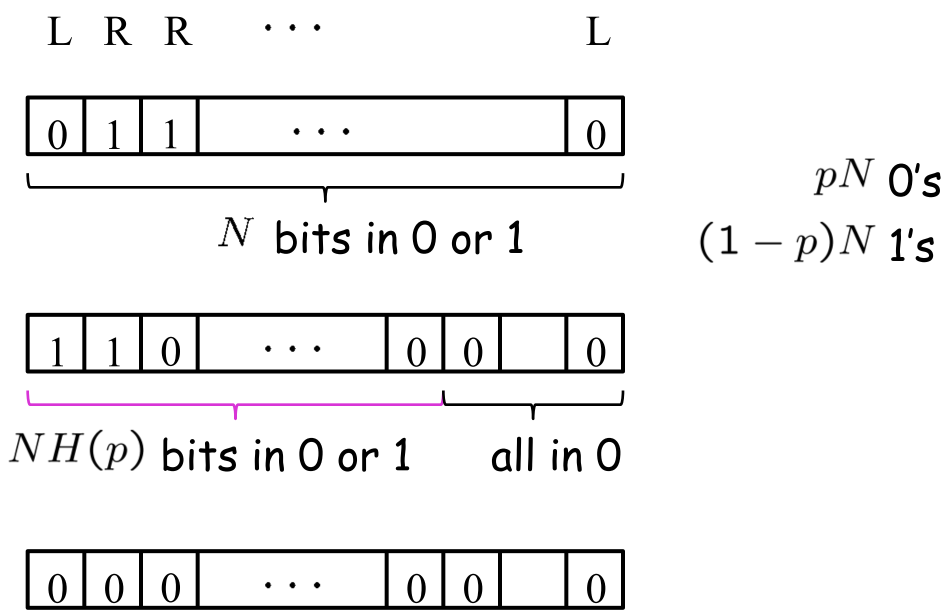
Initial memory state

Data compression

$N \to NH(p)$

Information erasure

$W\text{erasure} = kT \ln 2H(p)$



At the end of the day, the net work consumption is $W\text{erasure} - W\text{sze} \geq 0$

the equivalence between two entropies

The **cleverest** strategy the demon can take is the data compression, whose optimality is proven by Shannon.

↓

The (minimum) erasure work W_{erasure} coincides with W_{sze}.

↓

This fact augments the Landauer-Bennett argument on the equivalence between the two entropies.

No use of the optimality based on the free energy, $F = U - TS$.

- a consequence of the second law
- S is thermodynamic, rather than information theoretic.

A. Hosoya, KM, Y. Shikano, PRE 2011

Any insight into statistical mechanics as well?

Take the **Boltzmann distribution**, for example.

$$p_i \propto \exp(-E_i/kT)$$

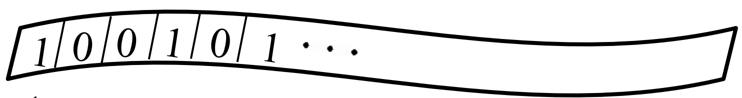
The Boltzmann distribution can be derived under the "principle of maximum (Shannon) entropy (PME)". But why should the Shannon entropy be maximised physically?

Lets derive it **operationally**, using the physics of erasure.

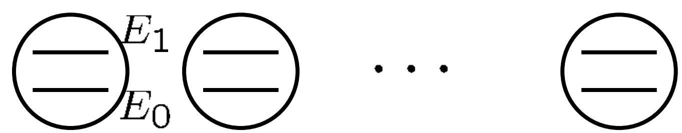
thermo-turing model

A system of a long tape and particles

(in contact with a heat bath of temperature T)



A tape storing a bit string



An ensemble of two-level particles

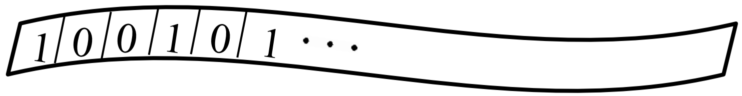
Each cell can be modelled as a memory with a single molecule

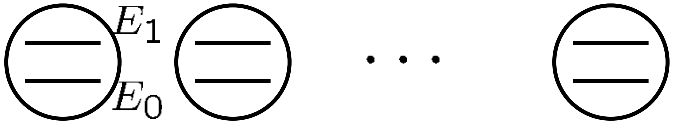
Demon manipulating the tape and the particles

Want to find p_0 and p_1 in the equilibrium condition.

A system of a long tape and particles

(in contact with a heat bath of temperature T)





Consider a cost function

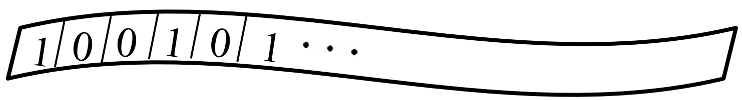
$$F = \varepsilon p_1 - kT \ln 2H(p_1)$$

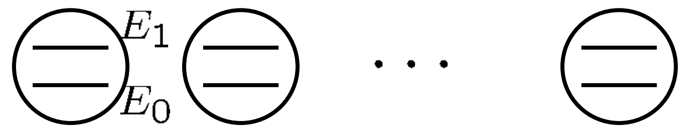
The equilibrium condition can be defined as $\Delta F = 0$, against random occurrences of NOTs (bit flips).

K. Maruyama

A system of a long tape and particles

(in contact with a heat bath of temperature T)





Consider a cost function

$$F = \varepsilon p_1 - kT \ln 2H(p_1)$$

Intuitively,

the 1st term : (average) energy to excite particles

the 2nd term : (average) energy to erase info on the tape

➡ An operational way to justify the Principle of Maximum Entropy

Boltzmann distribution

The net change in F due to a NOT is

$$\Delta F = \varepsilon(p_1 - p_0) - NkT \ln 2 \Delta H(p_0)$$

$$= \varepsilon(2p_0 - 1) - kT \ln 2 (1 - 2p_0) \frac{dH(p)}{dp}$$

$$\Delta F \equiv 0 \quad \Rightarrow \quad \frac{p_1}{p_0} = \exp(-\varepsilon/kT)$$ ☺

(No reference to a thin energy shell $E \sim E + \Delta E$)

Possible to generalise to multi-level systems

A. Hosoya, KM, Y. Shikano, in preparation

Fluctuation-Dissipation relation

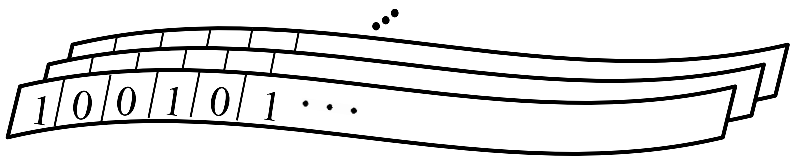
The thermo-Turing model may be applicable to nonequilibrium situations.

Now we attempt to obtain the Einstein relation $D = \sigma kT$

Diffusion (fluctuation) → D

mobility (dissipation) → σ

Consider an ensemble of tapes



For n-th cell, $\Delta F_n = u_n \neq 0$

An energy u_n from (unspecified) external source causes a bit flip to change the probability distribution, p_n.

Fluctuation-Dissipation relation

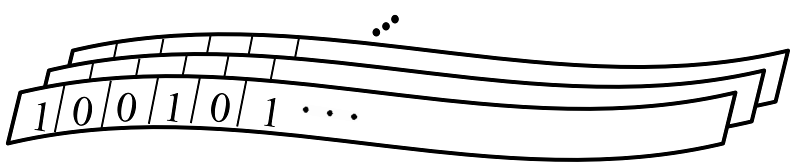
$$\Delta F_n = u_n \quad \Rightarrow \quad \frac{p_1}{p_0} = \exp(-(\varepsilon + u_n)/kT)$$

The average of work to make this change (u_n) happen (assuming $\sum_n u_n = 0$):

$$\langle W \rangle = \sum_n p_0^{(n)} u_n \propto -D/kT,$$

where $D := \sum u_n^2$.

Regarding $\langle W \rangle$ as mobility σ, this is essentially the Einstein relation .



A. Hosoya, KM, Y. Shikano, in preparation

Summary I

- Information is physical and is also subject to laws of physics

- The second law of thermodynamics is applicable to the 'sum' of entropies (thermodynamic + information)

- The second law is a cool meta-theory

- Thermo-Turing model could be useful in understanding thermodynamics/statistical mechanics in an operational way through info-thermo duality

KM, F. Nori, V. Vedral, Rev. Mod. Phys. 81, 1 (2009)
A. Hosoya, KM, Y. Shikano, Phys. Rev. E 84, 061117 (2011)
L. Brillouin, J. App. Phys. 22, 334 (1951)
E. T. Jaynes, Phys. Rev. 106, 620 (1957)