Relation between Work and Information

Guest lecture for Prof Mahn-Soo Choi's course on
Quantum Information Physics

Quantum Information &
Intelligent Energy Harvesting
Lab

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Imperial College
London

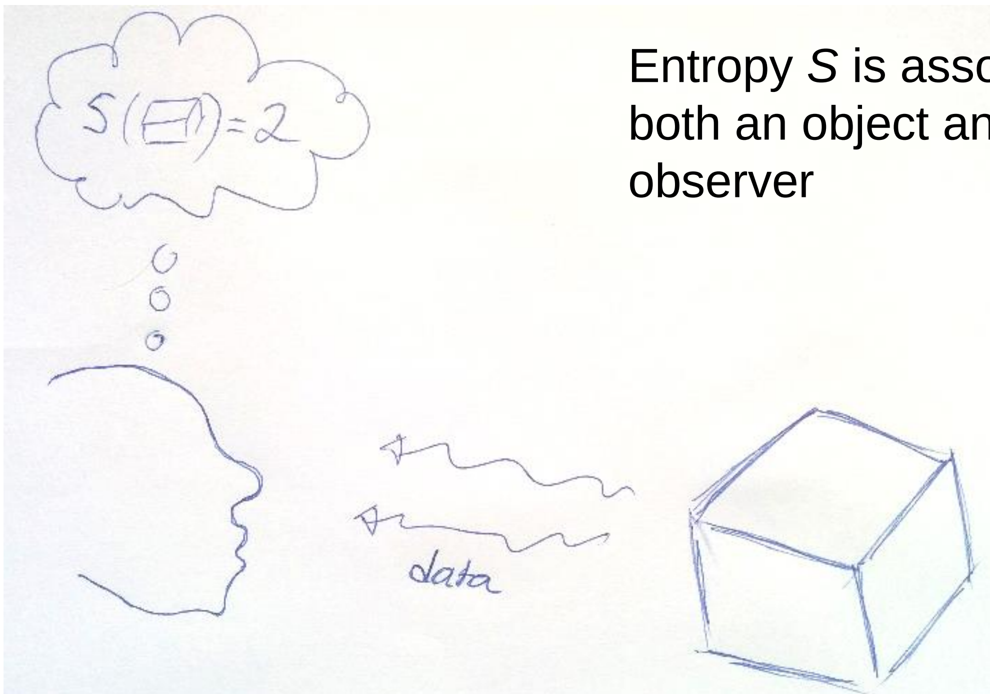
Oxford Physics

INSTITUTE FOR MATHEMATICAL SCIENCES LONDON
TRUTH BEAUTY

Oscar Dahlsten | dahlsten@sustc.edu.cn | Korea University May 2018

‘Relation between Work and Information’

- By having *information* about a system I mean the state we assign it is not uniformly random (entropy S not maximal).



Entropy *S* is associated with both an object and an observer

- ‘Work’ is essentially energy transferred into a battery.

- The ‘relation’ here is that ‘more information means more extractable work’.

Information and work extraction-why care?

1. The question is inherently interesting, e.g. how can we have equations with something subjective, S, and something objective, W, together?

2. Fundamental limits of work and heat immensely important for mankind.

   "The power densities of typical integrated circuits are approaching those of a light bulb filament ($\sim 100 \frac{W}{cm^2}$). Removal of the heat generated by an integrated circuit has become perhaps **the** *crucial constraint on the performance of modern electronics"**.

3. Energy is crucial to life, and as we shall discuss, the interplay between entropy and energy is too.

* *MIT Open course on nano-electronics.*

The relation between work and information : overview

1. Energy
2. Work (energy transferred)
3. Information
4. Inequality linking work and information:
   Free energy can only decrease.
5. Applications.
6. Our current research

Energy, how is it modelled

The Hamiltonian is the most important concept in modern physics
—Schroedinger.

Energy E is modelled as a real number, assigned to a state.
Classical mechanics: The state is x, p (or more such variables), and the assignment is done by a function $E = H(x, p)$ called the Hamiltonian. (Can also have probability distributions over x,p as our state).
Quantum mechanics: The state is a pure state ket $|\psi\rangle$ or a density matrix ρ. Average energy of state is assigned as

$$\langle H \rangle = \langle \psi | \hat{H} | \psi \rangle,$$

where $\hat{H}$ is the Hamiltonian operator.
In both cases **the Hamiltonian determines the time evolution**.

$$\frac{dx}{dt} = \frac{\partial H}{\partial p}; \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x};$$

$$\frac{d|\psi\rangle}{dt} = -i\hat{H}|\psi\rangle$$

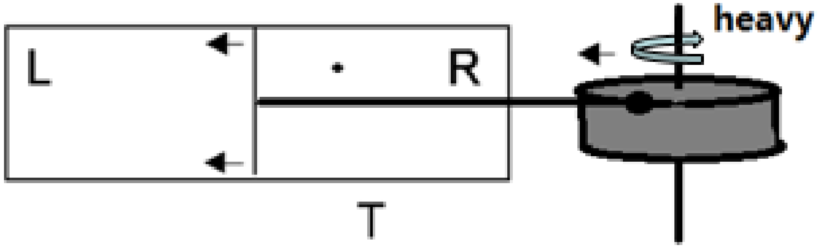
[Quantum and classical description can be unified: *Defining Hamiltonians beyond quantum theory, Branford, Garner and Dahlsten]*

Why/when is energy conserved

- In both classical and quantum mechanics, energy conserved under **time-independent** Hamiltonian.
- **Time-dependent** Hamiltonians can change energy. Then have implicit outside system too (such that total Hamiltonian time-independent).
- Example: Particle in a box, with

$$H(x,p) = \frac{p^2}{2m} + V(x,t),$$

where $V(x,t)$ is infinite at walls and 0 inside.



Implicit outside system

How to model energy transfer between systems?

Total energy of a closed system conserved under time evolution:

$$\Delta\langle H\rangle = 0,$$

but energy can move between subsystems A and B. We write

$$\langle H\rangle = \langle H_A\rangle + \langle H_B\rangle + \langle H_{int}\rangle,$$

where $\langle H_{A/B}\rangle$ must only depend on state of subsystem A/B, and H_{int} depends on both. Then

$$\Delta(\langle H_A\rangle + \langle H_B\rangle + \langle H_{int}\rangle) = 0$$

If $\langle H_{int}\rangle \approx 0$ at the points of evaluation, e.g. before and after an interaction takes place, energy of A and B individually well-defined.
Can now say energy was transferred between A and B:

$$\Delta(\langle H_A\rangle + \langle H_B\rangle) = 0.$$

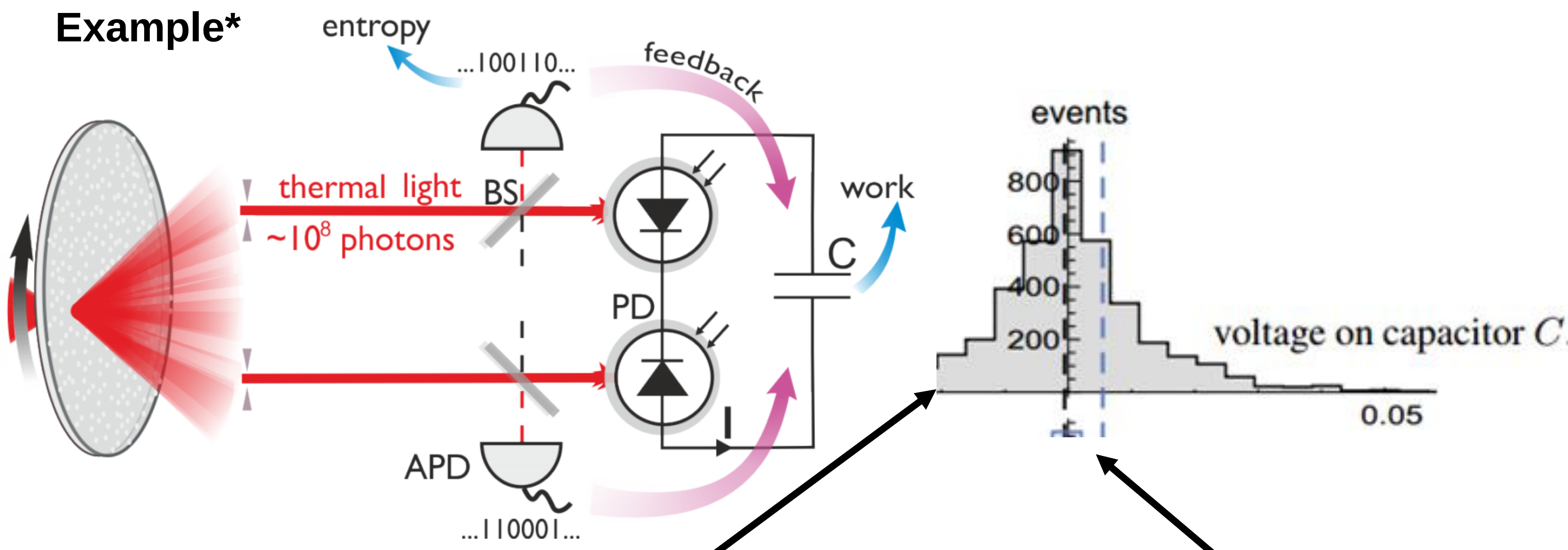
$$\Delta(\langle H_A\rangle) = -\Delta(\langle H_B\rangle)$$

.

What is work?

1. *Work out*: energy change of designated battery system.

Example*



2. Work has probability distrib., can focus on *average* work.

3. *Worst-case work* in any single shot also important.

*)Photonic Maxwells Demon,
Vidrighin, Dahlsten, Barbieri, Vedral, Walmsley, PRL 2016.

What is information

Entropy

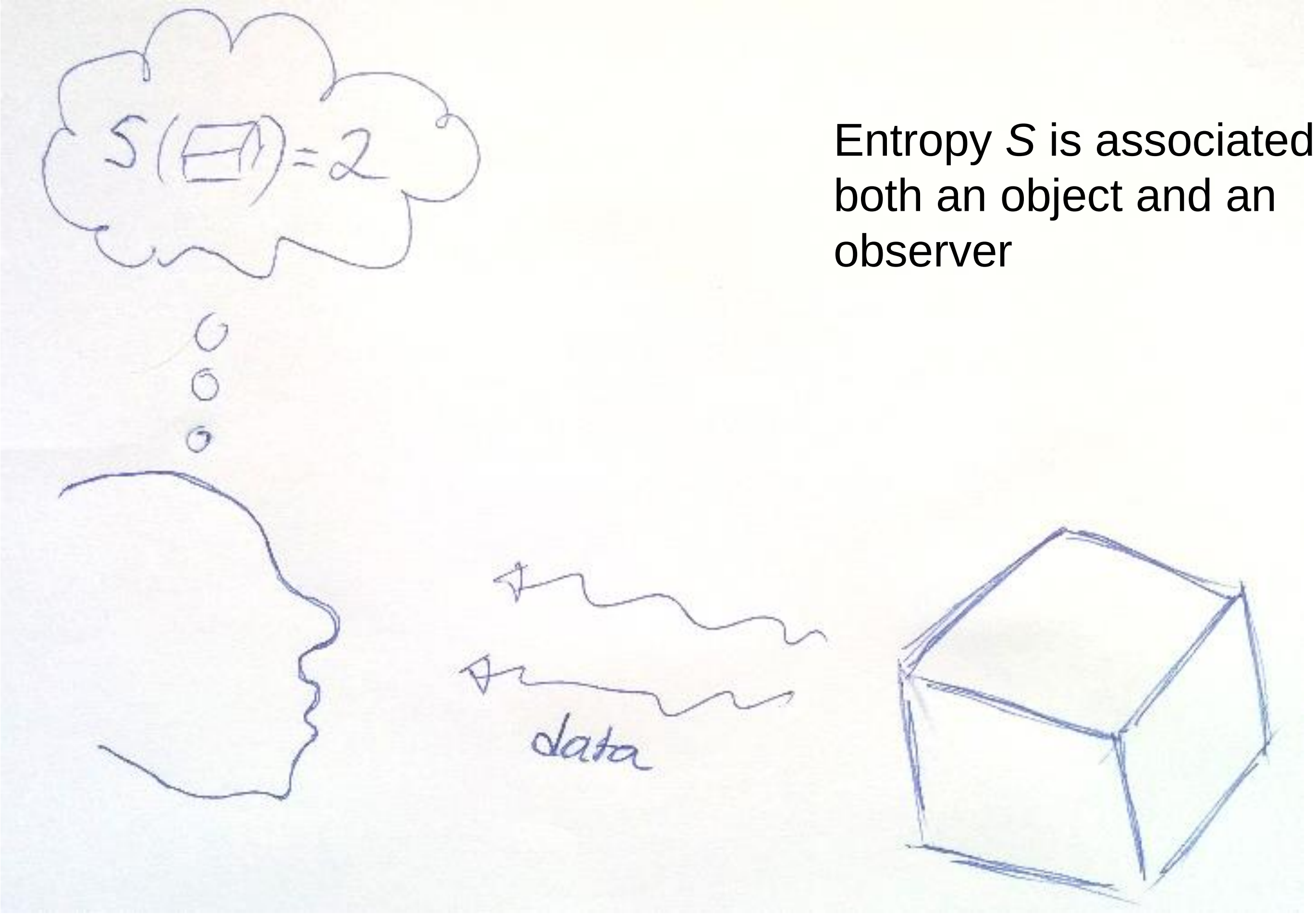
Our knowledge of a system is represented by mathematical objects called states:

1. Probability distribution $\vec{p}$ (Classical probability theory)
2. Density matrix ρ (Quantum theory)

The *entropy* is a real number assigned to each state representing essentially our uncertainty about the underlying microstate.

Conceptual: Probabilities are subjective

- Different people assign different probability distributions to the same events.
- Thus for a given function of probabilities —think *entropy*— they will assign different values.



Entropy *S* is associated with both an object and an observer

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function.
Shannon 1948

The great advance provided by information theory lies in the discovery that there is a unique, unambiguous criterion for the the "amount of uncertainty" represented by a discrete probability distribution [...]
Jaynes 1956

Information Entropy: how to quantify information

Examples of Entropy measures

Shannon entropy $S(X)_{\vec{p}} = \sum_i p_i \log(1/p_i)$.

Renyi entropies: $S_\alpha(X)_{\vec{p}} = \frac{1}{(1-\alpha)} \log \sum_i p_i^\alpha$.

Some special cases:

$$S_0(X)_{\vec{p}} = \sum_i \log p_i^0.$$

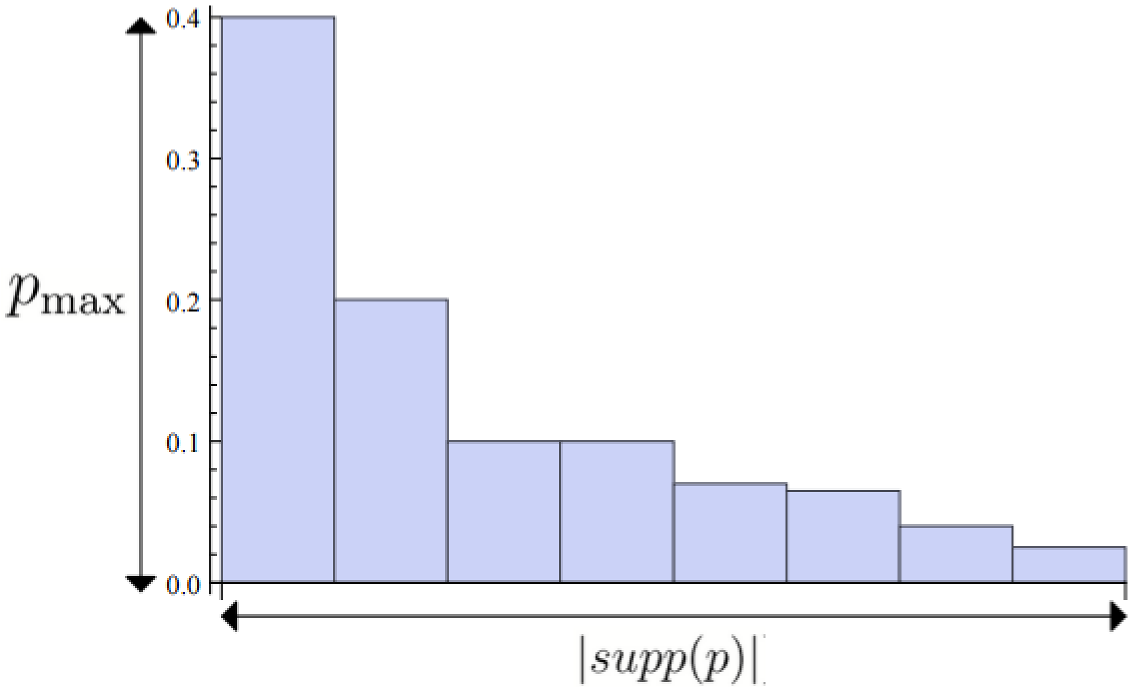
(Use L'Hopital's rule to see that)

$$S_{\alpha \to 1}(X)_{\vec{p}} = S(X)_{\vec{p}}.$$

$$S_2(X)_{\vec{p}} = -\log \sum_i p_i^2.$$

$$S_\infty(X)_{\vec{p}} = \log(p_{\max}).$$

Question, how large a memory?



In above example: 8 states $\rightarrow$ 2 bits

General Answer: $\sum_i p_i^0$ states $\rightarrow$ $\underline{\log \sum_i p_i^0 = S_0}$ bits.

Shannon entropy and large n iid limit

- It is common since Shannon to assume that the random process is independently and identically distributed (i.i.d): $\vec{p}_{tot} = \vec{p} \otimes \vec{p} \otimes ... \vec{p} = \vec{p}^{\otimes n}$.

- Such states allow for a neat type of information compression: as n gets large a large sequence of outcomes, such as 001100111..0 will very likely have about $np(0)$ 0's and $np(1)$ 1's. Thus effectively we have a probability distribution over such *typical* sequences alone.

- How many N_{typ} of them? Each has prob $p(0)^{np(0)}p(1)^{np(1)}$.

$$\Longrightarrow N_{typ}(p(0)^{np(0)}p(1)^{np(1)}) \leq 1$$

$$\Longleftrightarrow N_{typ} \leq 2^{\log(p(0)^{-np(0)}p(1)^{-np(1)})} = 2^{nH(\vec{p})}$$

- Here effectively there is a uniform distribution over 2^{nH} possible states, such that any Renyi entropy is effectively $\log(2^{nH(\vec{p})}) = nH(\vec{p})$.

Quantum entropy

Quantum theory state: density matrix ρ. Call eigenvalue spectrum $\vec{\lambda}$. In eigen-basis of ρ,

$$\rho = \sum_i \lambda_i |i\rangle\langle i|.$$

Von Neumann entropy

$$S(\rho) = \sum_i \lambda_i \log(1/\lambda_i).$$

Renyi entropy

$$S_\alpha(\rho) = \frac{1}{(1-\alpha)} \log \sum_i \lambda_i^\alpha.$$

Simplest interpretation: a measure of our ignorance of the underlying quantum state, 0 if state is pure.

Other entropic quantities

Note that the *Shannon entropy* is average surprise $(\log(1/p_i))$:
$S(\vec{p}) := \langle \log(1/p_i) \rangle_{\vec{p}}$

Relative entropy of two probability distributions:
$S(\vec{p}||\vec{q}) := \langle \log(p_i) - \log(q_i) \rangle_{\vec{p}}$
Average difference in surprise
of a given event bw two distributions p and q.

What is relation between work and information?

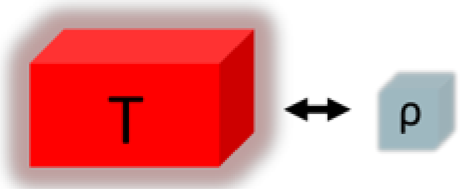
Inequality with energy and entropy

- Recall the thermal distribution over states labelled by i, with energies E_i:

$$\gamma = \sum_i \exp\left(-\beta E_i\right) |i\rangle\langle i| / Z,$$

where Z is a a normalisation factor.

- As things thermalise with a bath at temp T they get closer to this state: $\rho \to \gamma$



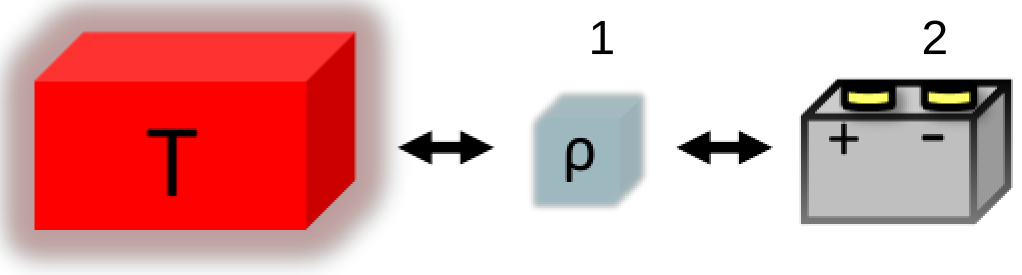
- It turns out this implies that $\langle H\rangle - TS$ can only decrease (then flattens out to the thermal state value which is $-kT \ln Z$).

A system thermalising with a heatbath at temp T will by assumption undergo a series of state changes all respecting

$$\Delta(\langle H\rangle - TS) \leq 0.$$

- Now we have an inequality with energy and entropy together.

Operational meaning of F=U-TS



Suppose battery (system 2) must have low entropy s.t. $F_2 = U_2 - TS_2 \approx U_2$, then

$$F_{12} \approx F_1 + F_2 \approx F_1 + U_2$$

As they interact with heat-bath alone, demand $\delta F_{12} \leq 0$ Thus

$$\delta F_2 \approx \delta U_2 \leq -\delta F_1 = \delta(U_1 - TS(\rho))$$

Let work $\langle W \rangle := \delta U_2$. Conclude

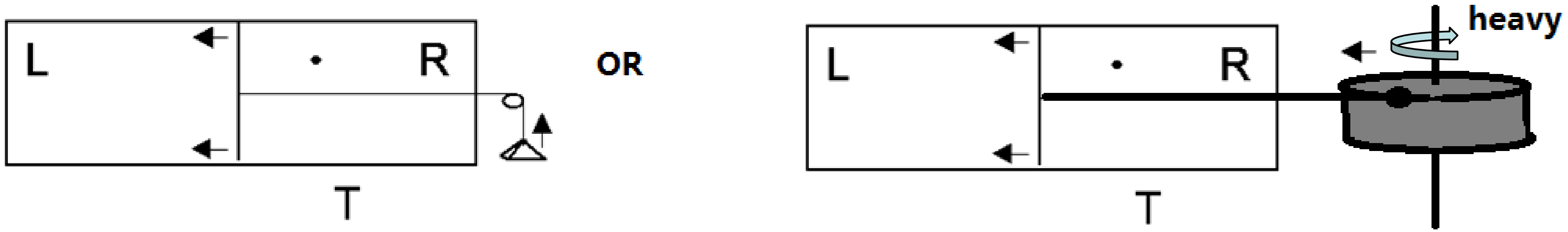
$$\langle W \rangle \leq \delta(U_1 - TS(\rho))$$

So F bounds the available (=free) work, and is therefore called the Free energy

Examples of relation between work and information

What is relation between work and information? Szilard engine illustrates

- There is a single particle in a box, and heat bath at temperature T.



- Daemon inserts divider in middle of the box, measures particle position, hooks up weight accordingly, then extracts work W isothermally.

$$\langle W \rangle \leq \delta F$$

Free energy has entropy in it

$$\delta F = \delta U - T\delta S = 0 - T\delta S = -Tk \ln 2,$$

where $\delta U = 0$ since by equipartition principle $U = (1/2)kT$ and T constant, and

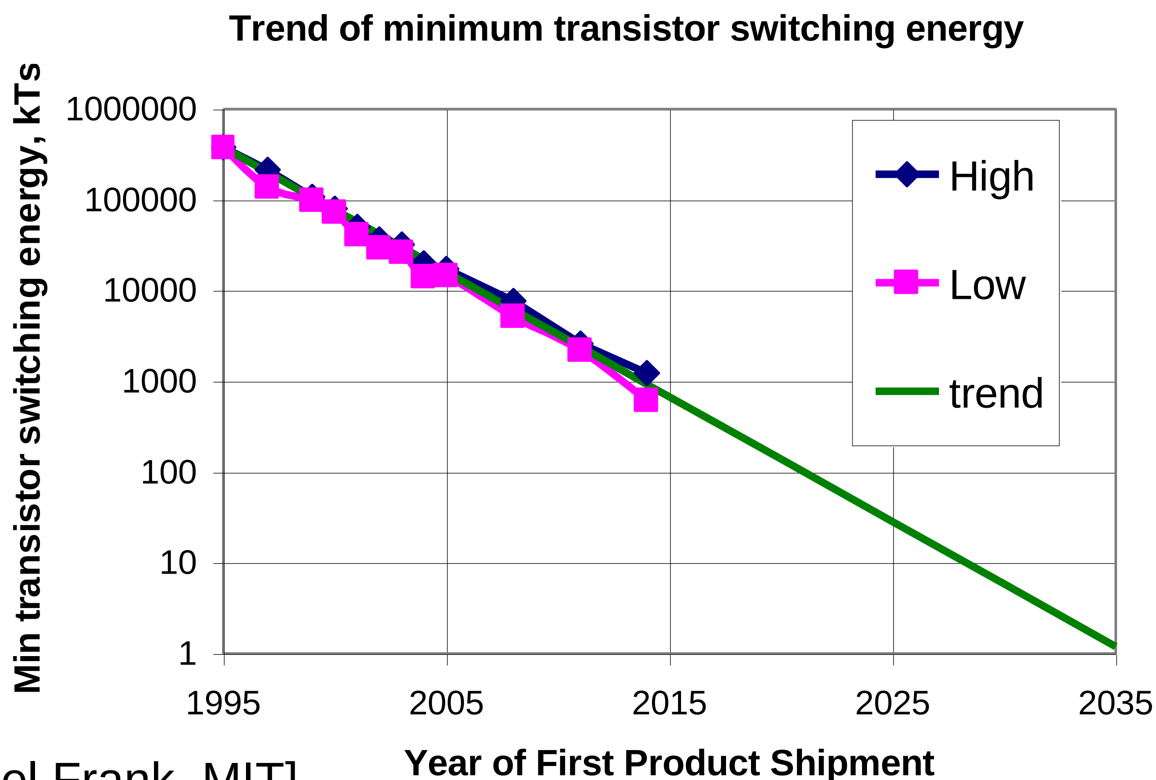
$\delta S = k(\ln 2d - \ln d) = k \ln 2$ where d is number of accessible phase space points.

Work value of one bit is kTln2

- We used up one bit (L vs R) to gain $\langle W \rangle = kT \ln 2$

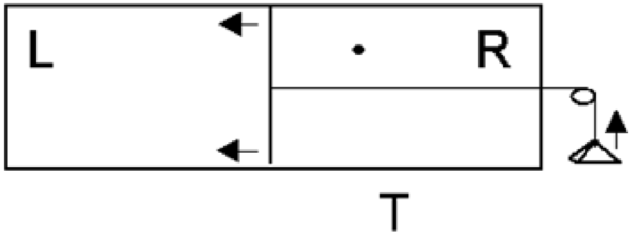
Landauer's principle for the reverse process.

- Suppose *reduce* a systems entropy by δS
- Must cost work at least $\delta(U - TS) = \delta(U) - T\delta(S)$.
- If set $\delta(U) = 0$, this means $W \geq T\delta(S)$.
- If we are resetting one bit, $(0,1) \to 1$ then $\delta S = k_B T \ln 2$. This is Landauers principle: **resetting one bit costs at least $kT \ln 2$ of work.**

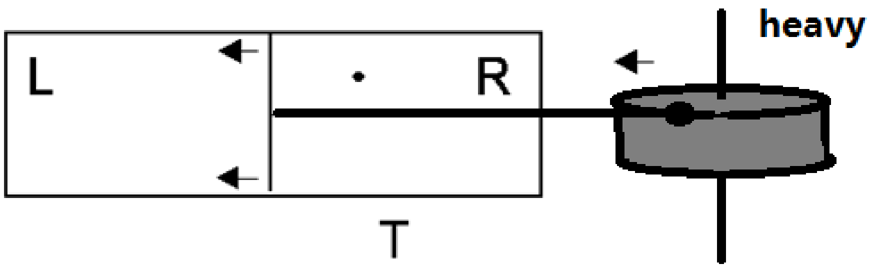


[Michael Frank, MIT]

Daemon (such as Szilard's) violates second law?



OR



Kelvin's version of second law.

> No process is possible in which the sole result is the absorption of heat from a reservoir and its complete conversion into work.

- After the Szilard engine work extraction the working medium is restored to its original state, seemingly violating Kelvin's law, **unless** there is a hidden work cost not accounted for.
- Bennett points out that the daemon may correlate its memory with the particle by a reversible interaction (CNOT gate) in principle at no energy cost. **Measurement does not need to cost work**.
- But, Bennett argues, the entropy of the memory is increased by 1 bit and by Landauer's principle it costs at least $kT \ln 2$ work to reset it. Bennett thus located the hidden work cost which saves Kelvin's law.

Life and Landauer's principle

1. Bacteria multiplying in solution use work, dissipating heat.

2. The creation of a new cell is a kind of reset:
   atoms must go into a restricted configuration.

BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 17, 1955

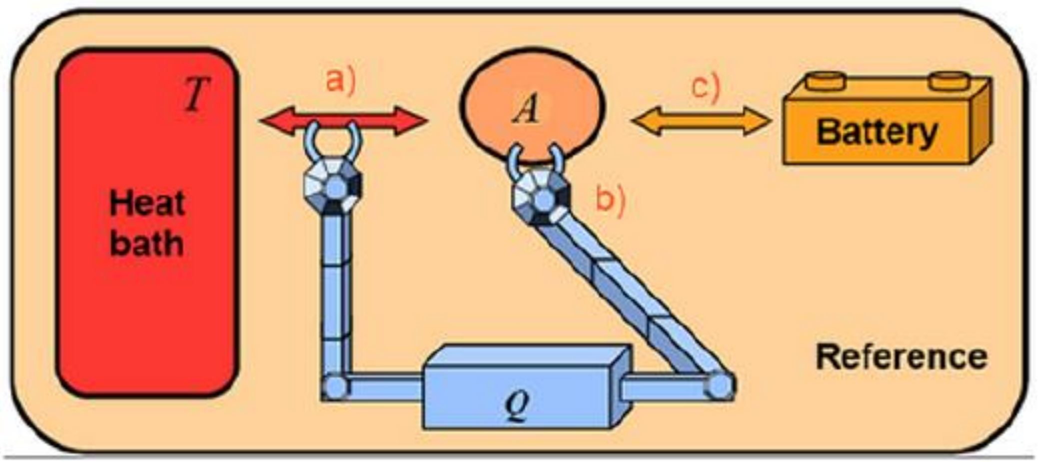
SOME ORDER-DISORDER CONSIDERATIONS IN LIVING SYSTEMS

HAROLD J. MOROWITZ*
LABORATORY OF TECHNICAL DEVELOPMENT, NATIONAL HEART INSTITUTE, NATIONAL INSTITUTES OF HEALTH, PUBLIC HEALTH SERVICE, U.S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE, BETHESDA, MD.

Order and disorder in biological systems are considered quantitatively in terms of information and entropy. After discussing the factors contributing to the information content of a living cell, a calculation is made of this parameter. The value for a typical bacterial cell is 4.6×10^{10} bits. This value is compared with an experimental value of the heat of growth and entropy production of *E. Coli*. A discussion of methods of improving the calculation is also

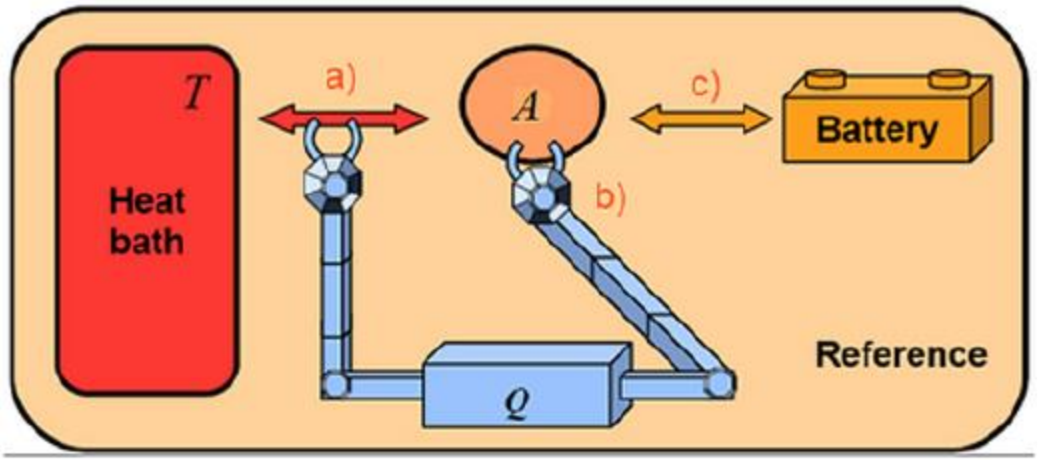
Thermodynamic Meaning of Negative entropy

- For von Neumann entropy S, conditional entropy given by: $S(A|Q) := S(AQ) - S(Q)$
- A simple example: $|\psi\rangle_{AQ} = |00\rangle + |11\rangle$, then $S(A|Q) = -1$.
- We want to interpret such negative entropy a' la Landauer/Szilard.
- **We include the observer's memory Q explicitly** in the description of the erasure of system A.



Del Rio et al, Nature 2011

Protocol erasing A with cost *S(A|Q)kTln2*:example



A simple example illustrates the general protocol:
$|\psi\rangle_{AQ} = |00\rangle + |11\rangle$, $S(A|Q) = -1$.

(i) Extract $W_{\mathrm{out}} = 2kT\ln 2$ work from **both** A and Q.

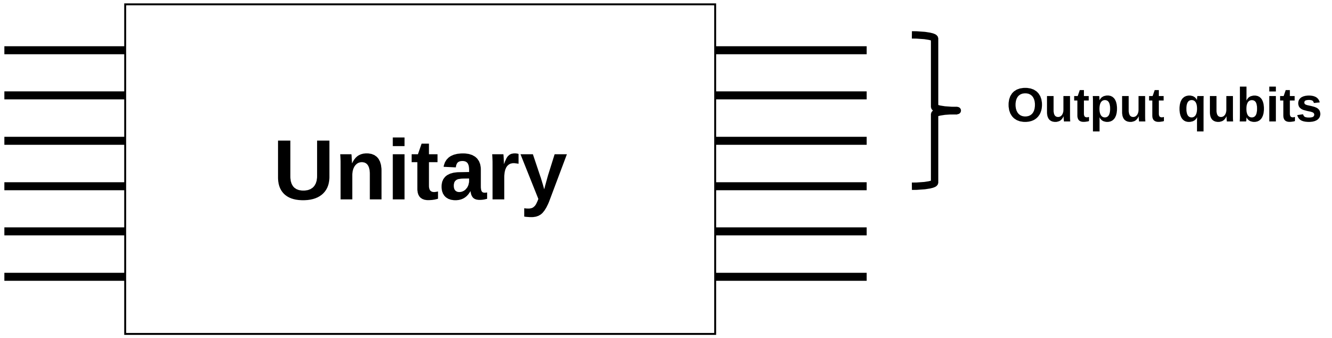
(ii) Reset A to $|0\rangle$ by using $W\mathrm{in} = kT\ln 2$ work.

Net result: A was reset to $|0\rangle$, reduced state on Q unchanged:

$W = W\mathrm{in} - W\mathrm{out} = -kT\ln 2 = S(A|Q)kT\ln 2.$

Application for cooling computers

- Extract work from correlations between the output qubits and the rest. Reduced state on output invariant.
- Consider circuit model computation, e.g. Shor‘s algorithm.



- Not all qubits are measured in the end to get the output.
- The energy extracted comes from the computer and its surroundings, so the computer is cooled.

Current state of art in work-information relation

1. Many mathematical expressions show
   "more information means more work"
   (though often hidden fundamental costs cancel the work output)

2. E.g. there are modifications of Jarzynski's equality to cases of feedback control*, optimal worst case work expressions (**) (single-shot statistical mechanics) and much more.

3. Several experiments consistent with the theory.

4. But arguably we are short of technological uses for "more information means more work". **We aim to find some.**

Reviews: Second Law-Like Inequalities with Quantum Relative Entropy: An Introduction T. Sagawa, arXiv:1202.0983v3;
The role of quantum information in thermodynamics—a topical review, Goold, Huber, Riera, del Rio, Skrzypczyk JPHYS A 2016
*Koski, Maisi, Sagawa, Pekola 1405.1272 PRL
**Dahlsten, Renner, Rieper, Vedral 0908.0424 NJP, del Rio, Aaberg, Renner, Dahlsten, Vedral 1009.1630 Nature Aaberg 1110.6121 NCOMMS, Horodecki, Oppenheim 1111.3834 NCOMMS and more recent work by eg Egloff, Gour, Faist, Yunger Halpern, Garner, ..

Summary & Outlook

- Work, information and their relation can be quantified.
- A key equation is
$$\langle W \rangle \leq \delta\langle H \rangle - T\delta S.$$
- "The extractable work is bounded by the free energy change".
- Different people can have different extractable work,(different S).

Now: making this useful for technology.

WE ARE HIRING
AT ALL LEVELS