Title: Laws of thermodynamics beyond the von Neumann regime

Date: Apr 29, 2013 04:00 PM

URL: http://pirsa.org/13040112

Abstract: A recent development in information theory is the generalisation of quantum Shannon information theory to the operationally motivated smooth entropy information theory, which originates in quantum cryptography research. In a series of papers the first steps have been taken towards creating a statistical mechanics based on smooth entropy information theory. This approach turns out to allow us to answer questions one might not have thought were possible in statistical mechanics, such as how much work one can extract in a given realisation, as a function of the failure-probability. This is in contrast to the standard approach which makes statements about average work. Here we formulate the laws of thermodynamics that this new approach gives rise to. We show in particular that the Second Law needs to be tightened. The new laws are motivated by our main quantitative result which states how much work one can extract or must invest in order to affect a given state change with a given probability of success. For systems composed of very large numbers of identical and uncorrelated subsystems, which we call the von Neumann regime, we recover the standard von Neumann entropy statements.

Joint work with Egloff, Renner and Vedral

http://arxiv.org/abs/1207.0434

The laws of thermodynamics beyond von Neumann regime *-recent results relating work and information*

Oscar Dahlsten
with Aaberg, Browne, del Rio, Egloff, Garner, Renner, Rieper and Vedral

oscar.dahlsten@physics.ox.ac.uk
oscar.dahlsten@nus.edu.sg

PI, Waterloo, April 2013

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIVERSITY OF OXFORD

NUS
National University of Singapore

Why study relation between work and information

1. Tension between subjectiveness of information and objectiveness of energy begs to be studied.

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. Landauers and Bennett's early arguments indeed guide nano-electronics research*.

4. Quantum information can bring its sophisticated understanding of *entropy* to the table.

5. Maybe information theory will get something back, recall that the von Neumann entropy came from thermodynamical considerations.

* *MIT Open course on nano-electronics.*

Why study relation between work and information

1. Tension between subjectiveness of information and objectiveness of energy begs to be studied.

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. Landauers and Bennett's early arguments indeed guide nano-electronics research*.

4. Quantum information can bring its sophisticated understanding of *entropy* to the table.

5. Maybe information theory will get something back, recall that the von Neumann entropy came from thermodynamical considerations.

* *MIT Open course on nano-electronics.*

Papers on 'smooth entropy thermodynamics'

This talk

	Title	Authors	arXiv	Journal
Part I	Inadequacy of von Neumann entropy for characterizing extractable work	Dahlsten, Renner, Rieper, Vedral	0908.0424	NJP
Bonus?	Thermodynamic meaning of negative entropy	del Rio, Aaberg, Renner, Dahlsten, Vedral	1009.1630	Nature
	Truly work-like work extraction	Aaberg	1110.6121	
	Fundamental limitations for quantum and nano thermodynamics	Horodecki, Oppenheim	1111.3834	
Part II	The laws of thermodynamics beyond the von Neumann Regime	Egloff, Dahlsten, Renner and Vedral	1207.0434	
	A Quantitative Landauer's Principle	Faist, Dupuis, Oppenheim, Renner	1211.1037	

Papers on 'smooth entropy thermodynamics'

This talk

	Title	Authors	arXiv	Journal
Part I	Inadequacy of von Neumann entropy for characterizing extractable work	Dahlsten, Renner, Rieper, Vedral	0908.0424	NJP
Bonus?	Thermodynamic meaning of negative entropy	del Rio, Aaberg, Renner, Dahlsten, Vedral	1009.1630	Nature
	Truly work-like work extraction	Aaberg	1110.6121	
	Fundamental limitations for quantum and nano thermodynamics	Horodecki, Oppenheim	1111.3834	
Part II	The laws of thermodynamics beyond the von Neumann Regime	Egloff, Dahlsten, Renner and Vedral	1207.0434	
	A Quantitative Landauer's Principle	Faist, Dupuis, Oppenheim, Renner	1211.1037	

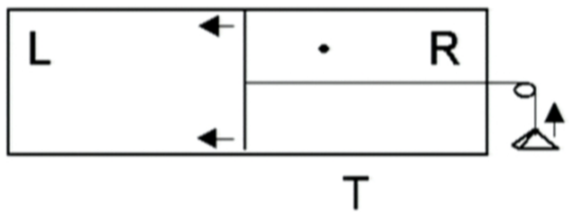
Papers on 'smooth entropy thermodynamics'

This talk

	Title	Authors	arXiv	Journal
Part I	Inadequacy of von Neumann entropy for characterizing extractable work	Dahlsten, Renner, Rieper, Vedral	0908.0424	NJP
Bonus?	Thermodynamic meaning of negative entropy	del Rio, Aaberg, Renner, Dahlsten, Vedral	1009.1630	Nature
	Truly work-like work extraction	Aaberg	1110.6121	
	Fundamental limitations for quantum and nano thermodynamics	Horodecki, Oppenheim	1111.3834	
Part II	The laws of thermodynamics beyond the von Neumann Regime	Egloff, Dahlsten, Renner and Vedral	1207.0434	
	A Quantitative Landauer's Principle	Faist, Dupuis, Oppenheim, Renner	1211.1037	

Starting point: Szilard engine

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



- The daemon (also called agent) inserts a divider in the middle of the box, measures the position of the divider and hooks up the weight accordingly.

- It can extract work isothermally:

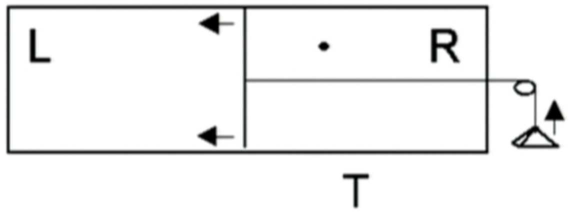
$$W = \int_V^{2V} p dV = \int_V^{2V} \frac{kT}{V} dV = kT \ln 2,$$

  where we used the ideal gas equation $pV = NkT$.

- Here we used up one bit to gain $kT \ln 2$ of work from heat bath. The inverse process is **Landauer's principle**: *it costs at least $kT \ln 2$ of work to erase (reset) one bit* (just change integration limits around).

Starting point: Szilard engine

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



- The daemon (also called agent) inserts a divider in the middle of the box, measures the position of the divider and hooks up the weight accordingly.

- It can extract work isothermally:

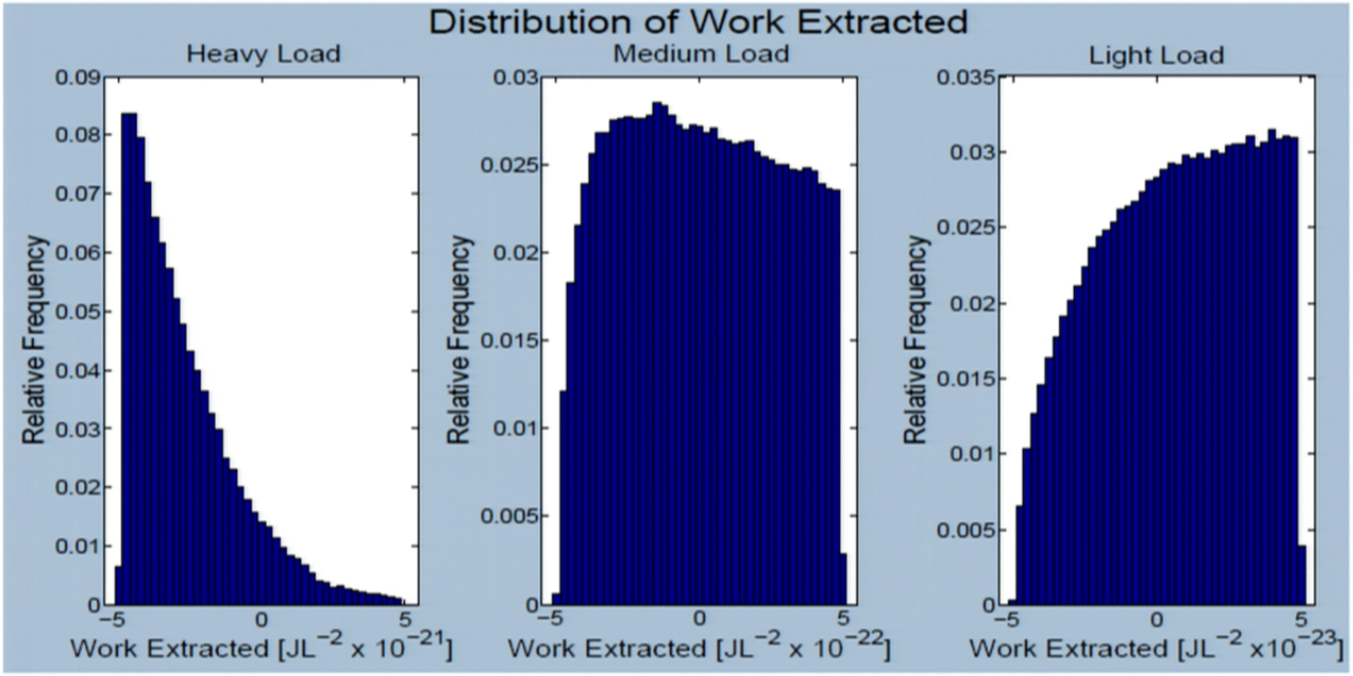
$$W = \int_V^{2V} p dV = \int_V^{2V} \frac{kT}{V} dV = kT \ln 2,$$

  where we used the ideal gas equation $pV = NkT$.

- Here we used up one bit to gain $kT \ln 2$ of work from heat bath. The inverse process is **Landauer's principle**: *it costs at least $kT \ln 2$ of work to erase (reset) one bit* (just change integration limits around).

Subtlety: get kTln2 in a single extraction?

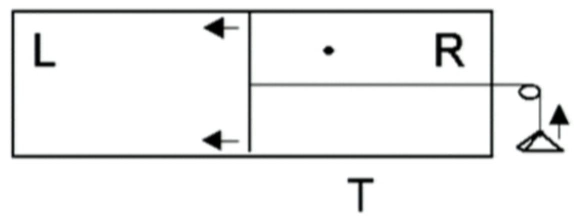
- If one models this engine naively like in the picture, one does not actually get $kT\ln 2$ in a single extraction, or on average (time or ensemble). *[C. Browne, EPSRC summer project, Oxford 2012][Discussions w. Aberg]]*



- Instead one should take the divider speed to be infinitely slow *independently* of the particle; Then $W = kT\ln 2$ with probability 1.

Starting point: Szilard engine

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



- The daemon (also called agent) inserts a divider in the middle of the box, measures the position of the divider and hooks up the weight accordingly.

- It can extract work isothermally:

$$W = \int_V^{2V} p dV = \int_V^{2V} \frac{kT}{V} dV = kT \ln 2,$$

where we used the ideal gas equation $pV = NkT$.

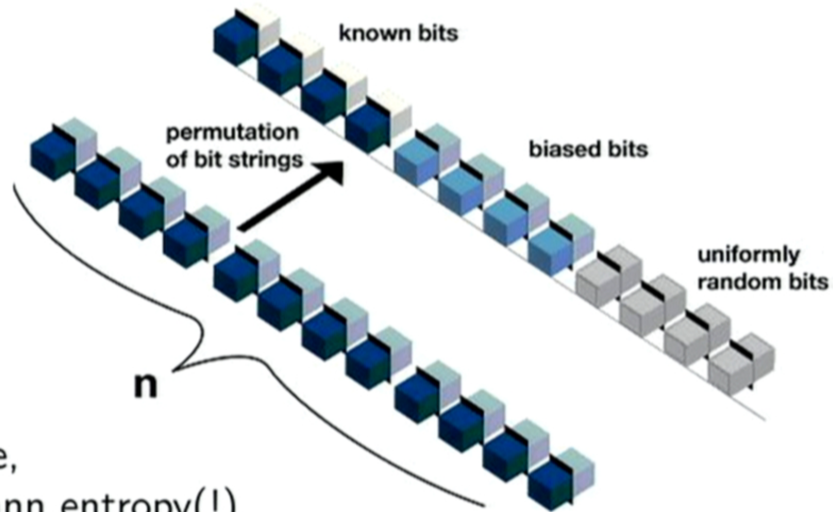
- Here we used up one bit to gain $kT \ln 2$ of work from heat bath. The inverse process is **Landauer's principle**: *it costs at least $kT \ln 2$ of work to erase (reset) one bit* (just change integration limits around).

Many Szilard engines: Bennett's information compression

- Say there are 2 Szilard boxes (n=2) associated with the state [p(LL) p(LR) p(RL) p(RR)]=[1/2 0 0 1/2].
- Can still extract $kT \ln 2$ of work by doing a CNOT such that we get [p(LL) p(LR) p(RL) p(RR)]=[1/2 0 1/2 0], and then extract work from the second box (known to be L).

[Bennett '82]

- More generally one can compress the randomness onto some bits and extract work from the rest. One may guess that optimally $W = kT \ln 2(\sharp\text{known bits})$, with $\sharp\text{known bits} = n - S(\rho)$.



- This turns out to be essentially true, but not for S being the von Neumann entropy(!).

Need to go beyond von Neumann entropy *S*

- The previous example of distribution (1/2 1/2 0 0) posed no challenge to the Shannon/von Neumann entropy $S_{vN} = \sum_i p_i \log_2 \frac{1}{p_i}$ being the appropriate entropy for thermodynamics (in quantum case $p_i \to \lambda_i$). The distribution is flat on some events and zero on the rest.
- For such distributions many entropy measures coincide. In particular the *Renyi entropies*,
$$S_\alpha := \frac{\log \sum_i p_i^\alpha}{1-\alpha}$$
all give $S = \log d$ for such distributions, where d is the number of events with non-zero probability.
- Note $S_{vN} = S_1$. Two other crucial ones are S_0 and S_∞.
Because $S_\infty \leq S_\alpha \leq S_0 \, \forall \alpha$, $S_0 := S_{\max}$, and $S_\infty := S_{\min}$.
- In quantum information theory Renner et. al. have shown that it is (modified versions of) $S_{\max}$ and $S_{\min}$ which capture the operational meanings one wants, and S_{vN} is only relevant when it happens to coincide with them.

Need to go beyond von Neumann entropy *S*

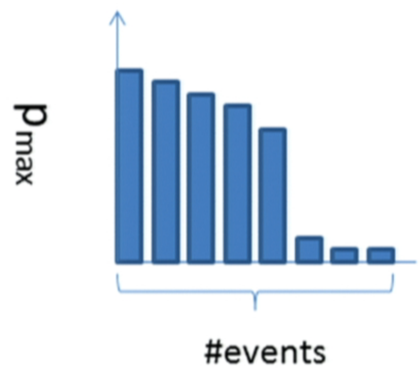
- The previous example of distribution (1/2 1/2 0 0) posed no challenge to the Shannon/von Neumann entropy $S_{vN} = \sum_i p_i \log_2 \frac{1}{p_i}$ being the appropriate entropy for thermodynamics (in quantum case $p_i \to \lambda_i$). The distribution is flat on some events and zero on the rest.
- For such distributions many entropy measures coincide. In particular the *Renyi entropies*,

$$S_\alpha := \frac{\log \sum_i p_i^\alpha}{1-\alpha}$$

all give $S = \log d$ for such distributions, where d is the number of events with non-zero probability.
- Note $S_{vN} = S_1$. Two other crucial ones are S_0 and S_∞.
Because $S_\infty \leq S_\alpha \leq S_0 \, \forall \alpha$, $S_0 := S_{\max}$, and $S_\infty := S_{\min}$.
- In quantum information theory Renner et. al. have shown that it is (modified versions of) $S_{\max}$ and $S_{\min}$ which capture the operational meanings one wants, and S_{vN} is only relevant when it happens to coincide with them.

Single-shot information theory, operational meaning of entropy

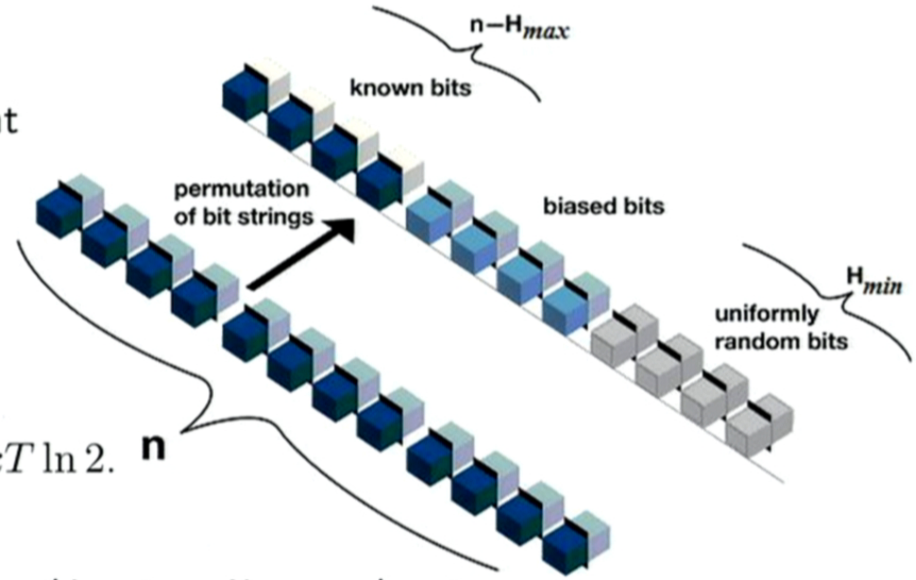
- One can show $S_0 := S_{\max} = \log(\sharp\text{events possible})$, and $S_\infty := S_{\min} = \log\left(\frac{1}{p_{\max}}\right)$
- An important operational meaning of S_0 is that it is the size a memory needs to have to reliably retain information.
- This is the size needed in *any given experiment*. It is therefore called a *single-shot* entropy.



Main results of Part I: not the von Neumann entropy

The work value of ones information (state) is determined by the smooth entropies alone.

1. Except with $p < \varepsilon$ the agent can be certain to extract $W = (n - H^{\varepsilon}_{\max})kT \ln 2$.
2. Except with $p < 2\varepsilon$ for an agent willing to risk failing to extract work $W \leq (n - H^{\varepsilon}_{\min} + \log(1/\varepsilon))kT \ln 2$.



So if we follow Feynman, then 'entropy' is not von Neumann's entropy

within Bennett's framework. Suppose we have a tape with N bits. We *define* the information, I, in the tape by the formula:

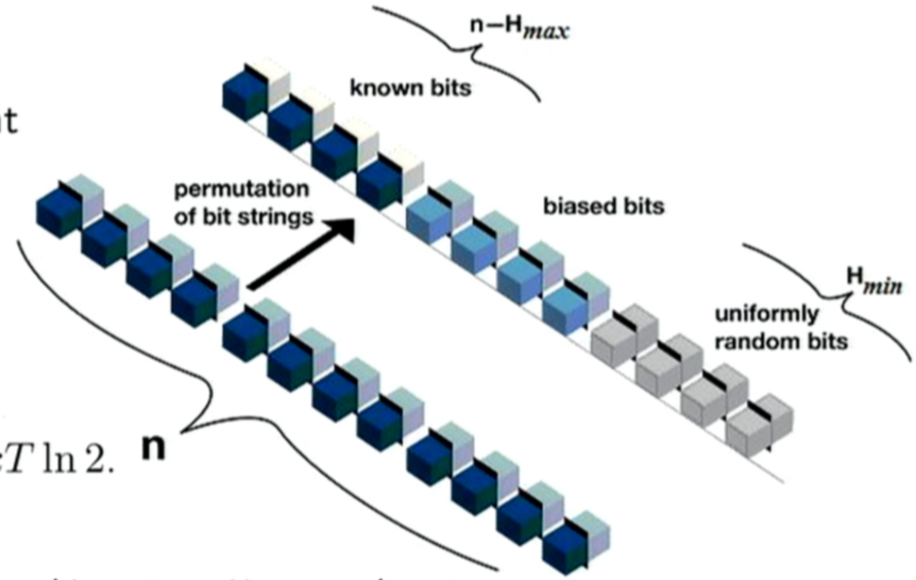
$$\textit{Fuel value of tape} = (N-I).kT \log 2. \qquad (5.12)$$

[Feynman '86]

Main results of Part I: not the von Neumann entropy

The work value of ones information (state) is determined by the smooth entropies alone.

1. Except with $p < \varepsilon$ the agent can be certain to extract $W = (n - H^{\varepsilon}_{\max})kT \ln 2$.
2. Except with $p < 2\varepsilon$ for an agent willing to risk failing to extract work $W \leq (n - H^{\varepsilon}_{\min} + \log(1/\varepsilon))kT \ln 2$.

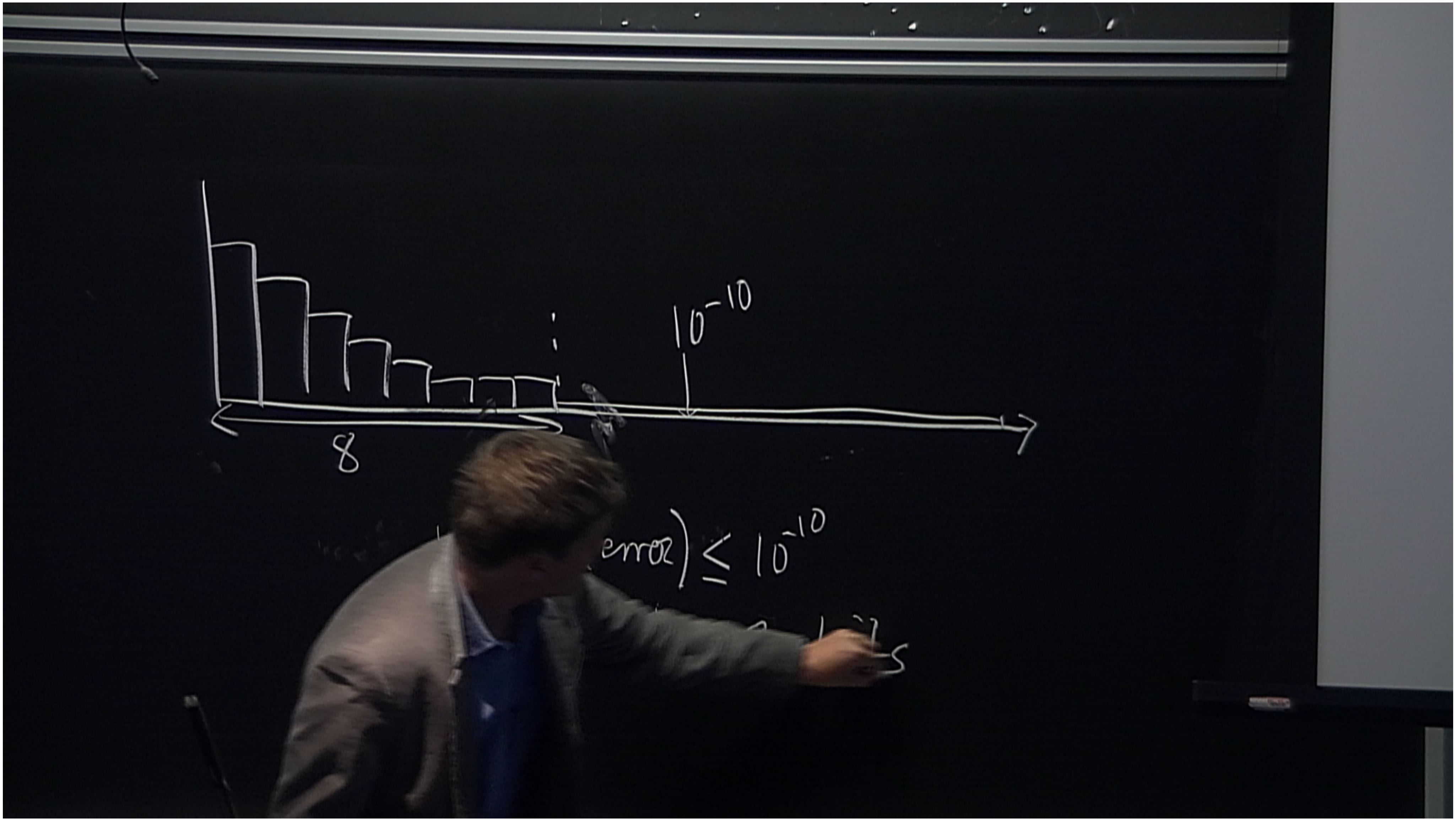


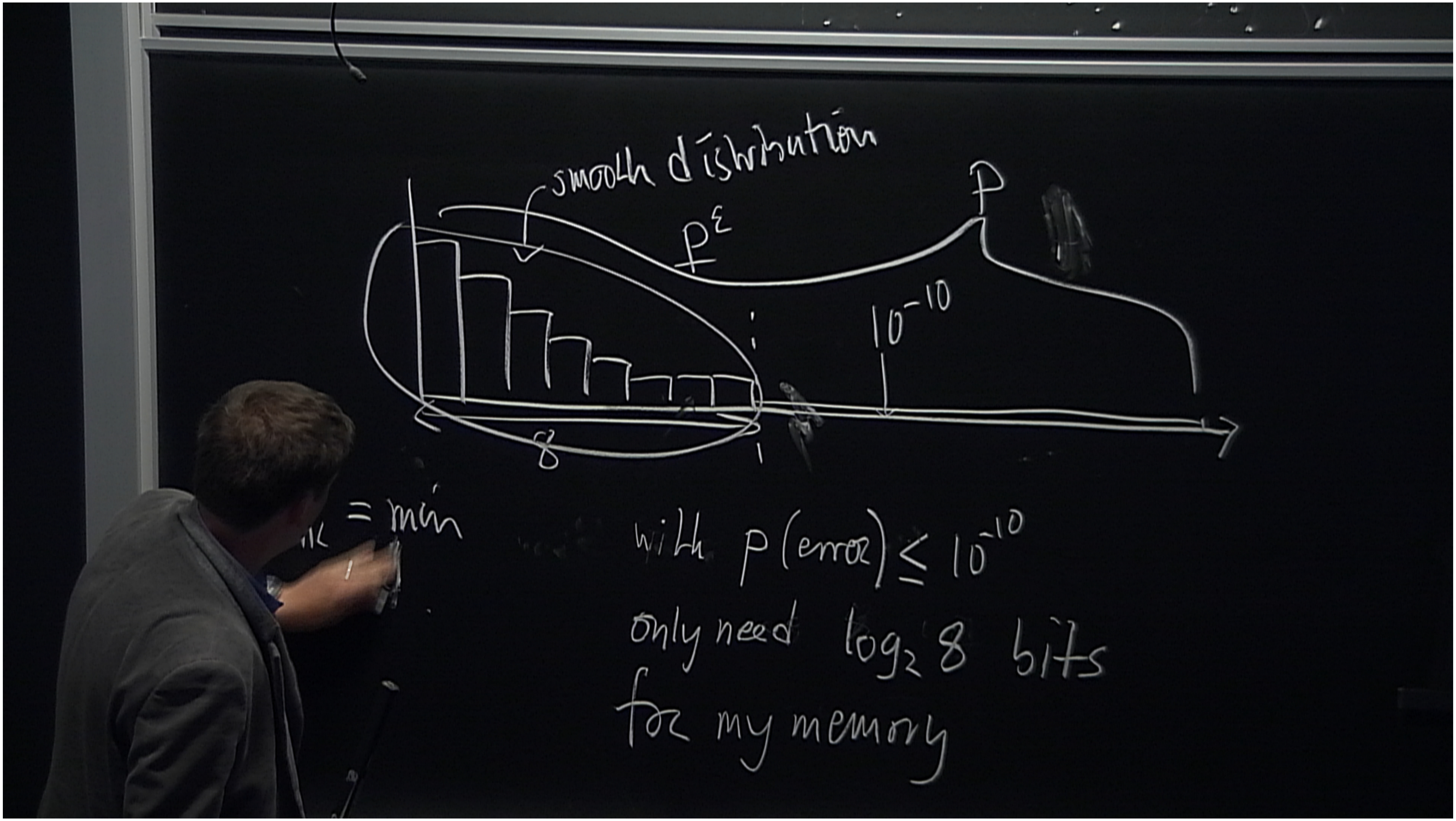
So if we follow Feynman, then 'entropy' is not von Neumann's entropy

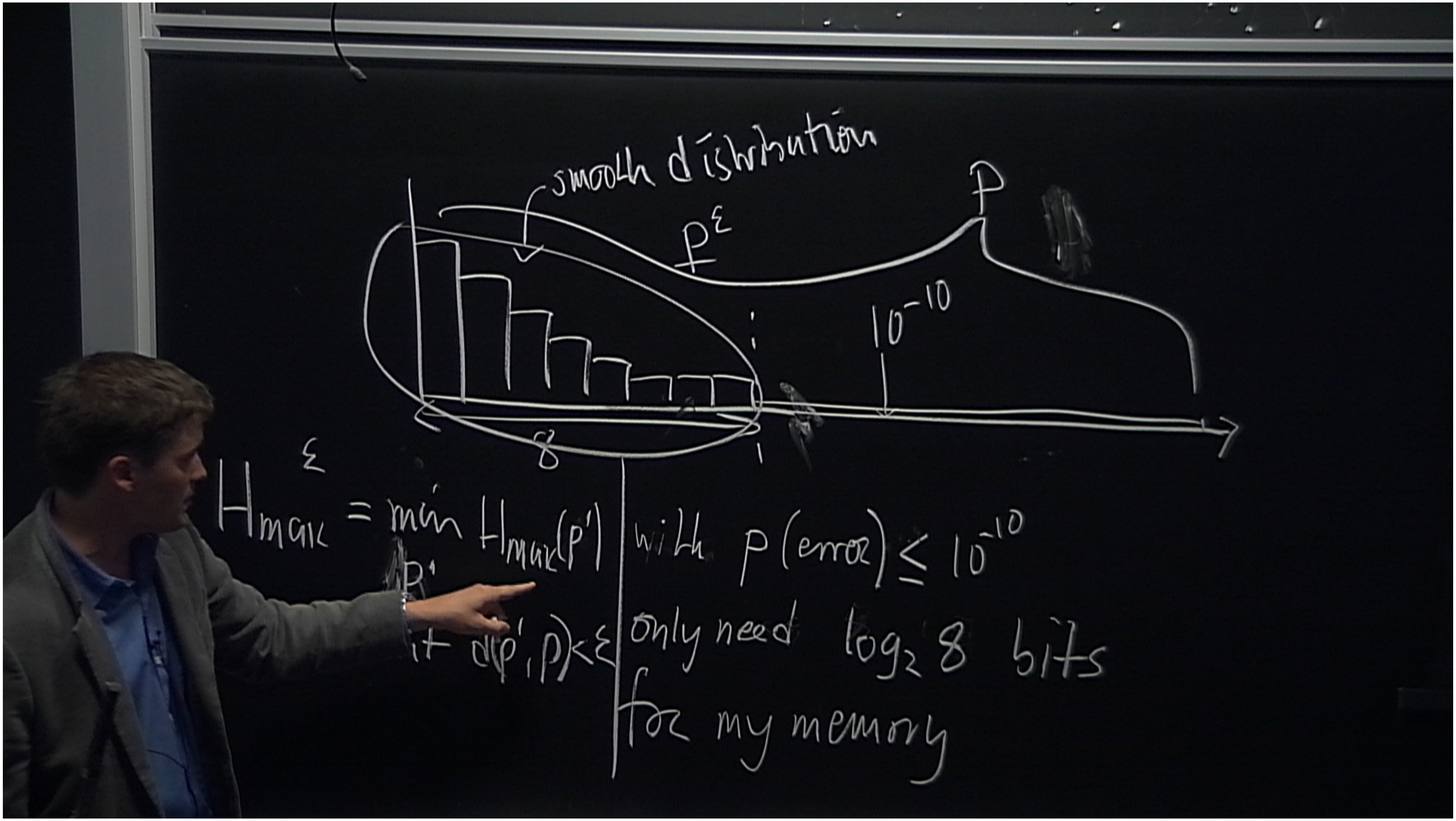
within Bennett's framework. Suppose we have a tape with N bits. We *define* the information, I, in the tape by the formula:

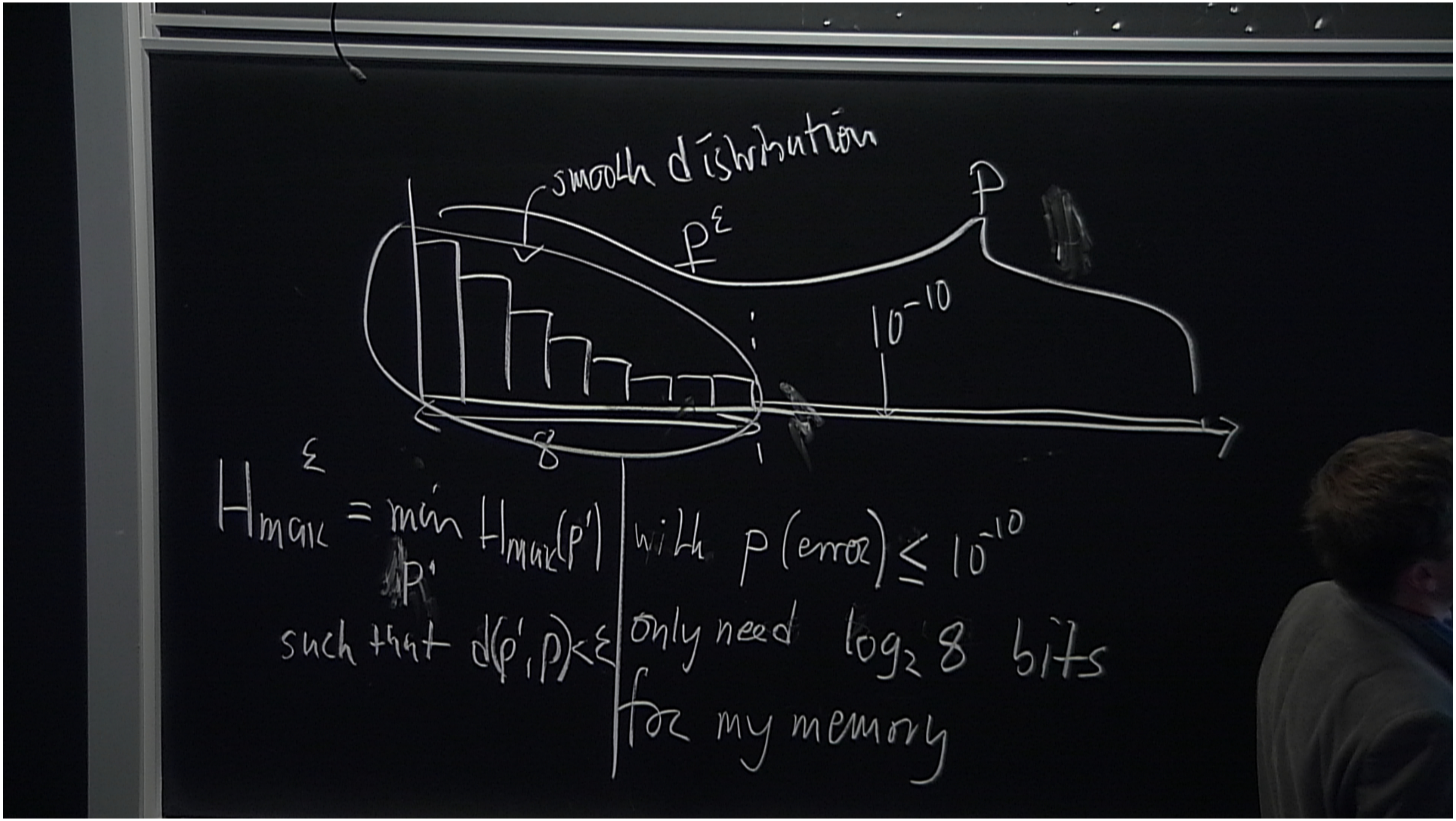
$$\textit{Fuel value of tape} = (N-I).kT \log 2. \qquad (5.12)$$

[Feynman '86]


10^{-10}
8
error) $\leq 10^{-10}$


smooth distribution
ρ
ρ^{ε}
10^{-10}
8
1
= min
with $p(\text{error}) \leq 10^{-10}$
only need $\log_2 8$ bits
for my memory


smooth distribution
P
P^{ε}
10^{-10}
8
$H^{\varepsilon}_{max} = \min_{P'} H_{max}(P')$ if $d(p',p) < \varepsilon$
with $p(\text{error}) \leq 10^{-10}$
only need $\log_2 8$ bits
for my memory


smooth distribution
p
p^{ε}
10^{-10}
8
$H_{max}^{\varepsilon} = \min_{p'} H_{max}(p')$
such that $d(p', p) < \varepsilon$
with $p(\text{error}) \leq 10^{-10}$
only need $\log_2 8$ bits
for my memory

Beyond the *what*?

By *von Neumann regime* we mean states which are of the form

$$\rho^{\otimes n}, \text{ with } n \to \infty.$$

We call it the von Neumann regime because

$$\lim_{n\to\infty,\varepsilon\to 0} \frac{S^{\varepsilon}_{\alpha}(\rho^{\otimes n})}{n} = S_{vN}(\rho) \;\; \forall\alpha,$$

i.e. in this regime the smooth entropies coincide
with the von Neumann entropy.

Beyond the *what*?

By *von Neumann regime* we mean states which are of the form

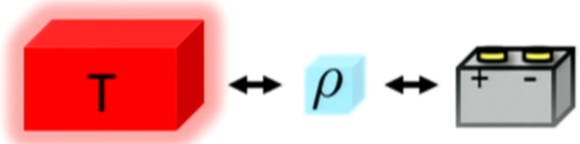
$$\rho^{\otimes n}, \text{ with } n \to \infty.$$

We call it the von Neumann regime because

$$\lim_{n\to\infty,\varepsilon\to 0} \frac{S_\alpha^\varepsilon(\rho^{\otimes n})}{n} = S_{vN}(\rho) \ \ \forall\alpha,$$

i.e. in this regime the smooth entropies coincide
with the von Neumann entropy.

A generalisation of the question



We saw in part I how we could gain work at the cost of randomising the state. Say now there is a more general state change $\rho \to \sigma$? How much work can one get?

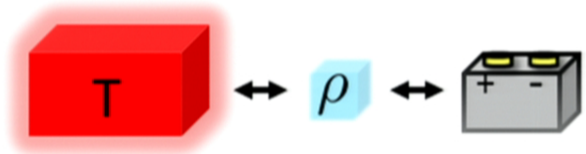
Subtlety: We are interested in the energy gained in any given realization, not on average. Why?
Because we care if $\Delta E(out) \geq \Delta E(threshold)$,
not whether $\langle \Delta E(out) \rangle \geq \Delta E(threshold)$.

We accordingly define $W(\rho \to \sigma)$
as the **maximal** $\Delta E(threshold)$ **that can be exceeded in any extraction given** ρ **and** σ.

Finally, it is natural to accept some probability ε of failing to achieve W. Then the work gain can be higher. The more general question now is:

$$W^{\varepsilon}(\rho \to \sigma) = ?$$

A generalisation of the question



We saw in part I how we could gain work at the cost of randomising the state. Say now there is a more general state change $\rho \to \sigma$? How much work can one get?

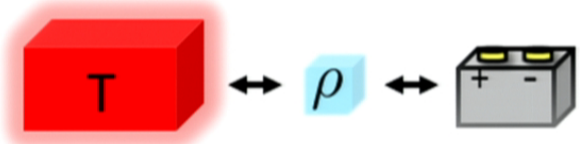
Subtlety: We are interested in the energy gained in any given realization, not on average. Why?
Because we care if $\Delta E(out) \geq \Delta E(threshold)$,
not whether $\langle \Delta E(out) \rangle \geq \Delta E(threshold)$.

We accordingly define $W(\rho \to \sigma)$ as the **maximal** $\Delta E(threshold)$ **that can be exceeded in any extraction given** ρ **and** σ.

Finally, it is natural to accept some probability ε of failing to achieve W. Then the work gain can be higher. The more general question now is:

$$W^{\varepsilon}(\rho \to \sigma) = ?$$

A generalisation of the question



We saw in part I how we could gain work at the cost of randomising the state. Say now there is a more general state change $\rho \to \sigma$? How much work can one get?

Subtlety: We are interested in the energy gained in any given realization, not on average. Why?
Because we care if $\Delta E(out) \geq \Delta E(threshold)$,
not whether $\langle \Delta E(out) \rangle \geq \Delta E(threshold)$.

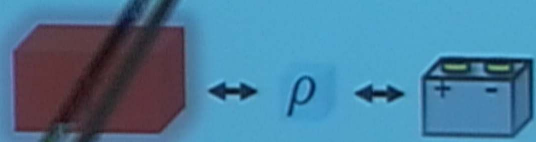
We accordingly define $W(\rho \to \sigma)$
as the **maximal** $\Delta E(threshold)$ **that can be exceeded in any extraction given** ρ **and** σ.

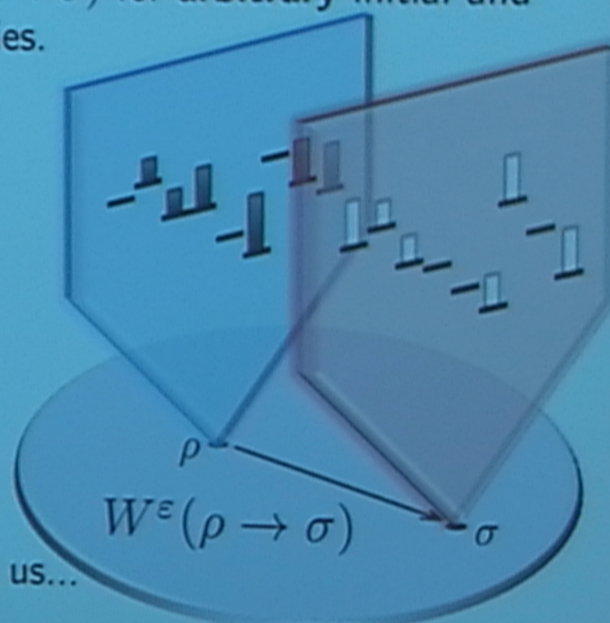
Finally, it is natural to accept some probability ε of failing to achieve W. Then the work gain can be higher. The more general question now is:

$$W^{\varepsilon}(\rho \to \sigma) =?$$

Radical Generalisations in calculating W^ε

- Aaberg and Horodecki & Oppenheim both go significantly beyond the paper discussed in part I, in particular they consider non-degenerate energy levels.
- Both papers arrive (in a priori different models(!)) at the same more general expression for how much work one can gain by taking some initial state to a thermal state on the same energy levels (a relative renyi-entropy).
- In this paper we go further, calculating $W^\varepsilon(\rho \to \sigma)$ for **arbitrary** initial and final energy spectra and occupation probabilities.

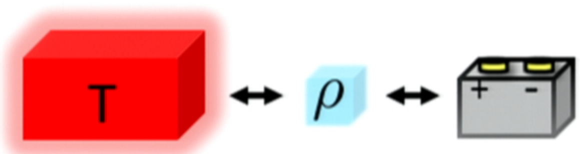


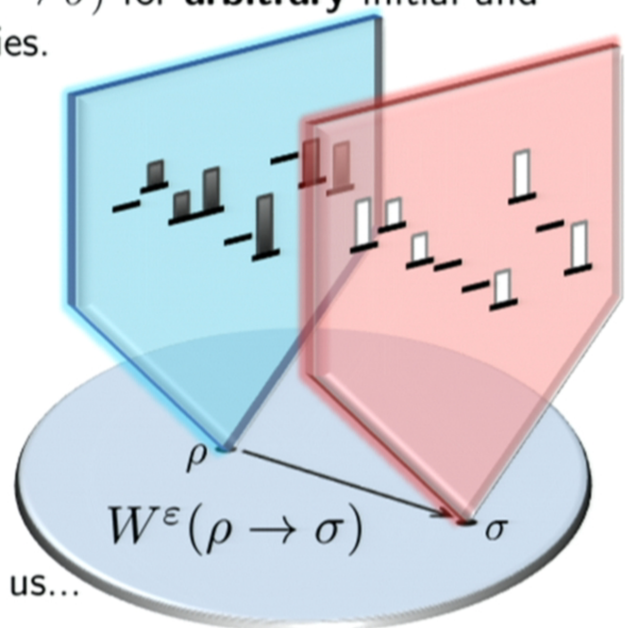


- To get an intuition for the work extraction scenario within which we derive $W^\varepsilon(\rho \to \sigma)$, consider a very simple example that has guided us...

Radical Generalisations in calculating W$^{\varepsilon}$

- Aaberg and Horodecki & Oppenheim both go significantly beyond the paper discussed in part I, in particular they consider non-degenerate energy levels.
- Both papers arrive (in a priori different models(!)) at the same more general expression for how much work one can gain by taking some initial state to a thermal state on the same energy levels (a relative renyi-entropy).
- In this paper we go further, calculating $W^{\varepsilon}(\rho \to \sigma)$ for **arbitrary** initial and final energy spectra and occupation probabilities.

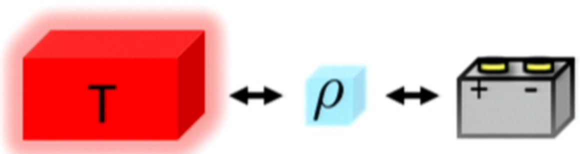


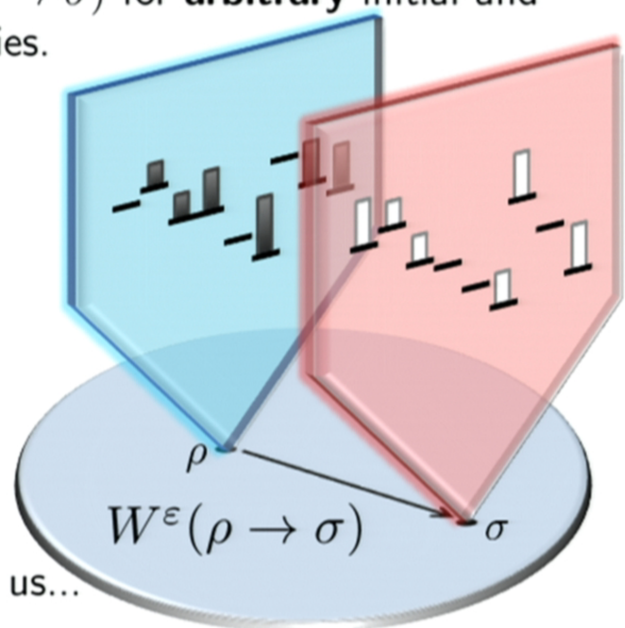


- To get an intuition for the work extraction scenario within which we derive $W^{\varepsilon}(\rho \to \sigma)$, consider a very simple example that has guided us...

Radical Generalisations in calculating W$^{\varepsilon}$

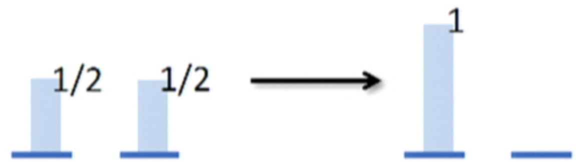
- Aaberg and Horodecki & Oppenheim both go significantly beyond the paper discussed in part I, in particular they consider non-degenerate energy levels.
- Both papers arrive (in a priori different models(!)) at the same more general expression for how much work one can gain by taking some initial state to a thermal state on the same energy levels (a relative renyi-entropy).
- In this paper we go further, calculating $W^{\varepsilon}(\rho \to \sigma)$ for **arbitrary** initial and final energy spectra and occupation probabilities.





- To get an intuition for the work extraction scenario within which we derive $W^{\varepsilon}(\rho \to \sigma)$, consider a very simple example that has guided us...

Simple example of work extraction scenario



[Alicki, 3 Horodeckis , 2004.]

- How? Raise the second level to infinity quasistatically and isothermally, so that the system is constantly in its thermal state

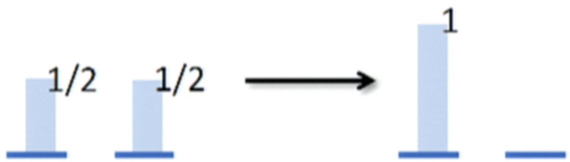
$$\rho_T = \sum_i \frac{\exp(-\beta E_i)}{Z} |e_i\rangle\langle e_i|.$$

- **Iff level is occupied when raised by δE_2, it costs δE_2 work.**
(consistent with: $dU = d\sum_i p_i E_i = \sum_i (dp_i) E_i + p_i (dE_i) := dQ + dW$)

$$\langle dW \rangle = p(E_1)0 + p(E_2)dE_2 = p(E_2)dE_2 = \frac{\exp(-\beta E_2)}{Z} dE_2.$$

$\Rightarrow W = \int_0^\infty \frac{\exp(-\beta E_2)}{Z} dE_2 = kT \ln 2$, (integration requires some few lines).

Simple example of work extraction scenario



[Alicki, 3 Horodeckis , 2004.]

- How? Raise the second level to infinity quasistatically and isothermally, so that the system is constantly in its thermal state

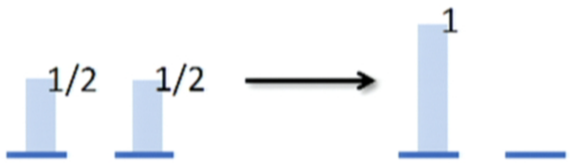
$$\rho_T = \sum_i \frac{\exp(-\beta E_i)}{Z} |e_i\rangle\langle e_i|.$$

- **Iff level is occupied when raised by δE_2, it costs δE_2 work.**
(consistent with: $dU = d\sum_i p_i E_i = \sum_i (dp_i)E_i + p_i(dE_i) := dQ + dW$)

$$\langle dW\rangle = p(E_1)0 + p(E_2)dE_2 = p(E_2)dE_2 = \frac{\exp(-\beta E_2)}{Z} dE_2.$$

$\Rightarrow W = \int_0^\infty \frac{\exp(-\beta E_2)}{Z} dE_2 = kT\ln 2$, (integration requires some few lines).

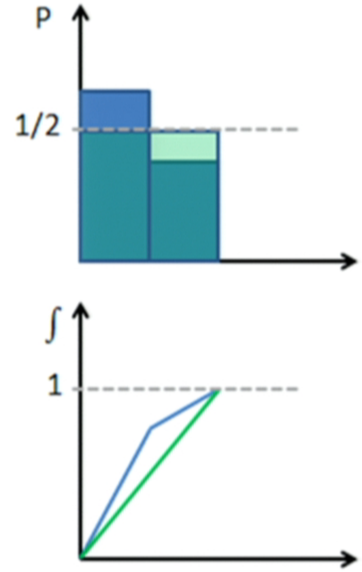
Simple example continued, how ε enters

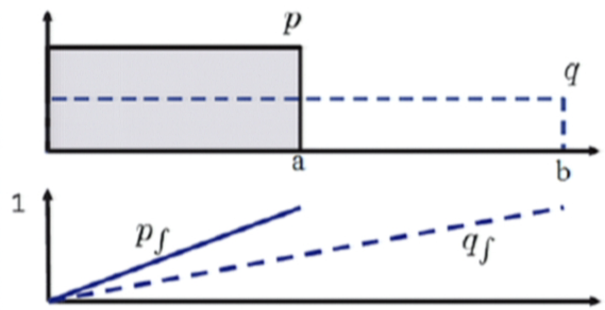


- Once the second level is very high the probability of being in level 2 is very low, call it ε.
- Except with probability ε it then suffices to input W^ε of work in order to take the initial state to the final.
- If you calculate this you will find $W^\varepsilon = -kT(\ln 2 + \ln(1-\varepsilon)) \leq W^0$.
- We can calculate W^ε for **arbitrary** initial and final energy spectra and occupation probabilities. To give the expression need two definitions...

Definition 1: Relative Mixedness *M*

- We define a measure of how much more random a distribution is than another.
- It is related to **majorisation**.
  A distribution majorises another if its integral (when both distributions are in descending order) always upper bounds the other.



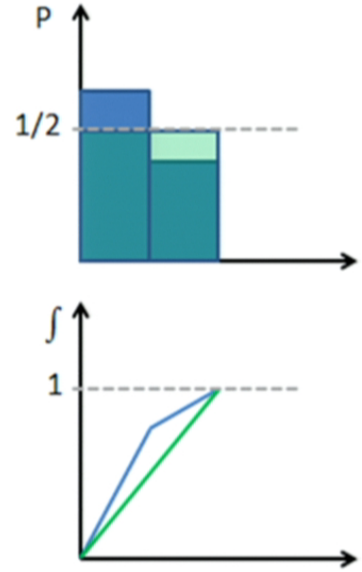


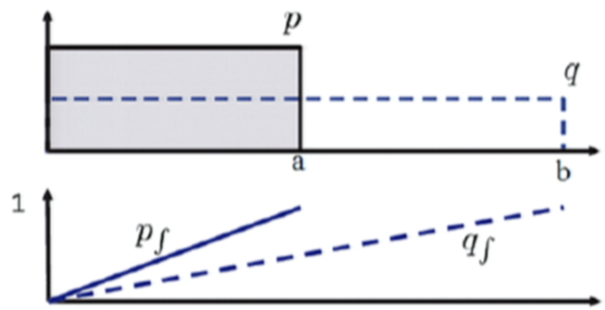
- The relative mixedness of two states, $M(\rho||\sigma)$, we define as the maximum stretching factor by which one can stretch the spectrum of one such that it still upper bounds the other.
  In simple case to the right $M = b/a$.
- M puts a number to *how much a distribution majorises another.*

Definition 1: Relative Mixedness *M*

- We define a measure of how much more random a distribution is than another.

- It is related to **majorisation**.
  A distribution majorises another if its integral (when both distributions are in descending order) always upper bounds the other.





- The relative mixedness of two states, $M(\rho||\sigma)$, we define as the maximum stretching factor by which one can stretch the spectrum of one such that it still upper bounds the other.
  In simple case to the right $M = b/a$.

- M puts a number to *how much a distribution majorises another.*

Understanding why one might think of Gibb's rescaling

- Consider the evolution of the energy level occupation probabilities under an interaction with the heat bath. There are good reasons to model this as a stochastic matrix (cols sum to 1), e.g. in the simplest case

$$\begin{pmatrix} p_0 \\ p_1 \end{pmatrix} \to \begin{pmatrix} p_0' \\ p_1' \end{pmatrix} = \begin{pmatrix} p_{(0\to 0)} & p_{(1\to 0)} \\ p_{(0\to 1)} & p_{(1\to 1)} \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}$$

- The stochastic matrix should leave thermal state invariant.
- We want to use a convenient class of stochastic matrices: those which leave the uniform distribution invariant (bistochastic matrices). But the thermal state is **not** in general uniform...
- We may however modify the thermal distribution to a uniform one by splitting events to make it uniform e.g.

$$\begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix} \to \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} ; \begin{pmatrix} p_{(0\to 0)} & p_{(1\to 0)} \\ p_{(0\to 1)} & p_{(1\to 1)} \end{pmatrix} \to \begin{pmatrix} p_{(0\to 0)}/2 & p_{(0\to 0)}/2 & p_{(1\to 0)}/2 \\ p_{(0\to 1)}/2 & p_{(0\to 1)}/2 & p_{(1\to 0)}/2 \\ p_{(0\to 1)} & p_{(0\to 1)} & p_{(1\to 1)} \end{pmatrix}$$

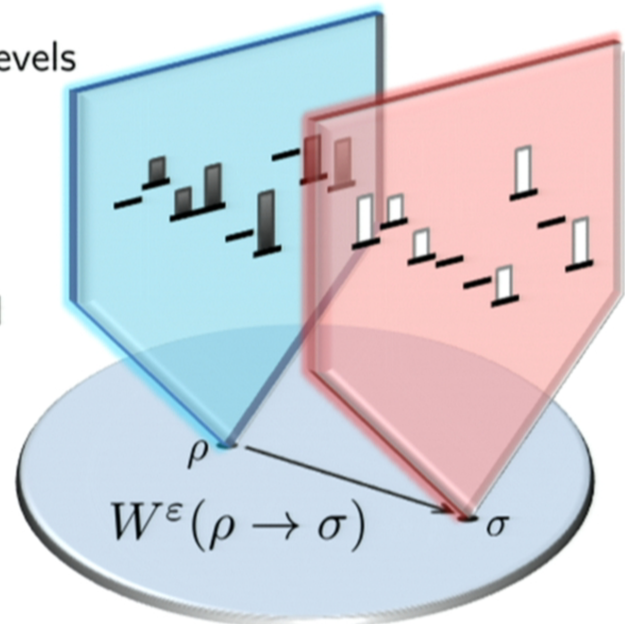
This splitting is the Gibb's rescaling. New matrix is bi-stochastic

Main result: expression for optimal W^{ε}

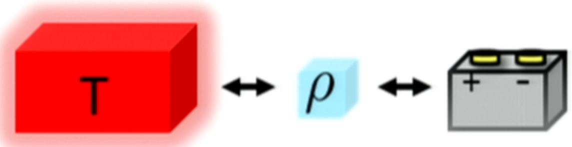
We prove that for arbitrary initial and final occupation probabilities and energy spectra, **the optimal $W^{\varepsilon}(\rho \to \sigma)$ is given by the relative mixedness M of the Gibbs-rescaled distributions**.

Expression by Aaberg and Horodecki & Oppenheim is important special case.

Part I expression recovered for degenerate levels
$W^{\varepsilon}(\rho \to \sigma_T) = (n - S^{\varepsilon}_{\mathsf{max}}(\rho))kT \ln 2.$
(Aaberg already showed it
is special case of their expression)



The expression also reduces to the standard
$F = -kT \ln \frac{Z_f}{Z_i}$) in appropriate limit.
In general it can be very different though.



Another result: Quantitative Second law needs tightening

- A standard expression is that

$$\Delta\left(S_{vN} - \beta U\right) \geq 0, \quad (1)$$

where U is expected internal energy of a system interacting with a heat-bath with inverse temperature β.

- In our model this is necessary but not *sufficient* to guarantee that an evolution is possible. Instead a state change $\rho \to \rho'$ due to a thermalisation with a heat-bath at temperature T is possible if and only if

$$W^0(\rho \to \rho') \geq 0. \quad (2)$$

- There are processes that respect Eq.1 but violate Eq.2. **We show such evolutions could be used to deterministically extract work from a single heat bath in a cycle.**

[Inspired also by Ruch, Meade et. al. (1970's)]

Another result: Quantitative Second law needs tightening

- A standard expression is that

$$\Delta\left(S_{vN} - \beta U\right) \geq 0, \tag{1}$$

where U is expected internal energy of a system interacting with a heat-bath with inverse temperature β.

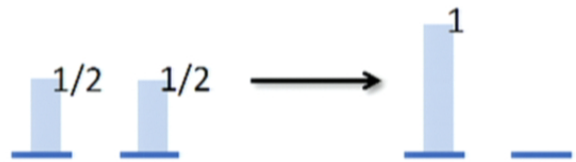
- In our model this is necessary but not *sufficient* to guarantee that an evolution is possible. Instead a state change $\rho \rightarrow \rho'$ due to a thermalisation with a heat-bath at temperature T is possible if and only if

$$W^0(\rho \rightarrow \rho') \geq 0. \tag{2}$$

- There are processes that respect Eq.1 but violate Eq.2. **We show such evolutions could be used to deterministically extract work from a single heat bath in a cycle.**

[Inspired also by Ruch, Meade et. al. (1970's)]

Simple example of work extraction scenario



[Alicki, 3 Horodeckis , 2004.]

- How? Raise the second level to infinity quasistatically and isothermally, so that the system is constantly in its thermal state

$$\rho_T = \sum_i \frac{\exp(-\beta E_i)}{Z} |e_i\rangle\langle e_i|.$$

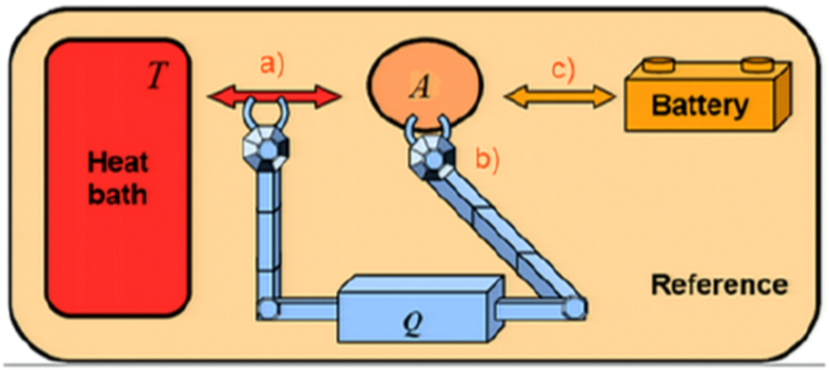
- **Iff level is occupied when raised by δE_2, it costs δE_2 work.**
(consistent with: $dU = d\sum_i p_i E_i = \sum_i (dp_i) E_i + p_i (dE_i) := dQ + dW$)

$$\langle dW \rangle = p(E_1)0 + p(E_2)dE_2 = p(E_2)dE_2 = \frac{\exp(-\beta E_2)}{Z} dE_2.$$

$\Rightarrow W = \int_0^\infty \frac{\exp(-\beta E_2)}{Z} dE_2 = kT \ln 2$, (integration requires some few lines).

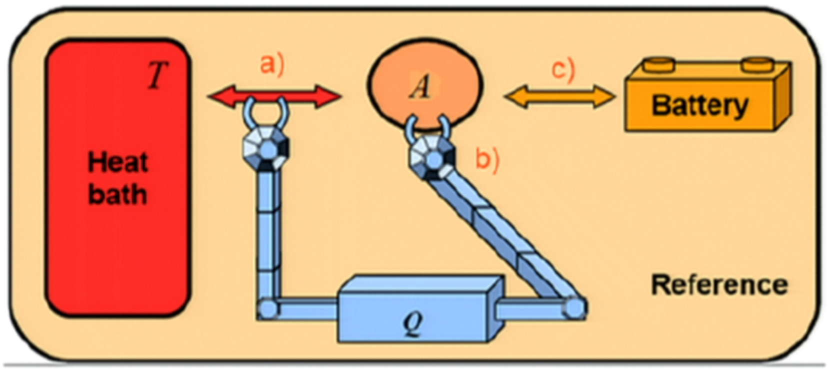
Negative entropy

- For von Neumann entropy S, conditional entropy given by: $S(A|Q) := S(AQ) - S(Q)$ (am supressing single-shot aspect in this section)
- A simple example: $|\psi\rangle_{SQ} = |00\rangle + |11\rangle$, 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.

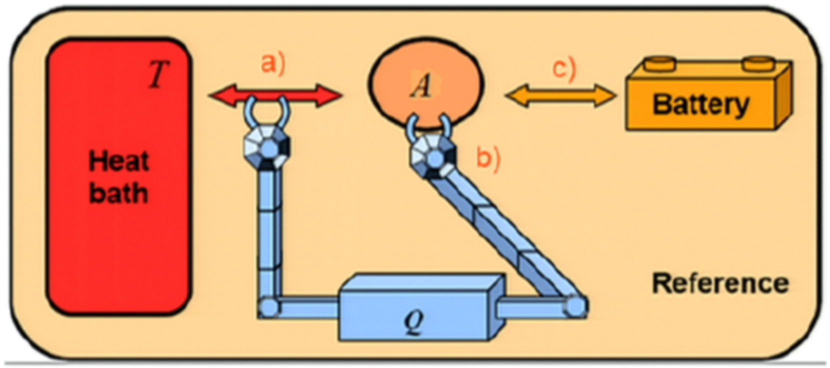


Negative entropy

- For von Neumann entropy S, conditional entropy given by: $S(A|Q) := S(AQ) - S(Q)$ (am supressing single-shot aspect in this section)
- A simple example: $|\psi\rangle_{SQ} = |00\rangle + |11\rangle$, 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.



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



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

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

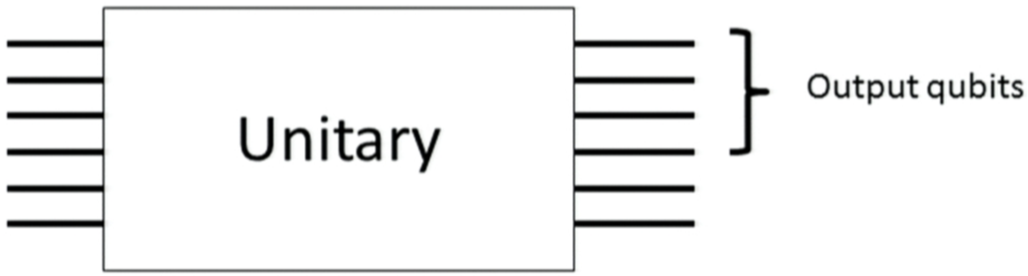
(ii) Reset A to $|0\rangle$ by using $W\text{in} = kT\ln 2$ work.

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

$W = W\text{in} - W\text{out} = -kTln2 = S(A|Q)kTln2.$

Application for cooling computers

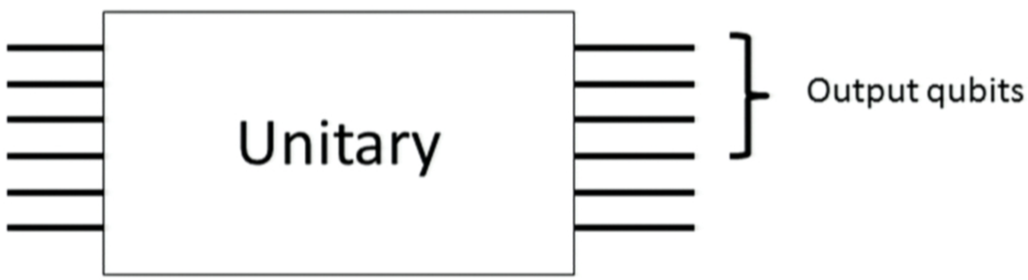
- Extract work from correlations between the output qubits and the rest. Reduced state on output invariant.
- Consider circuit model computation, e.g. Shor's algorithm.



- 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.

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.

Thank you

Joint work *with Aaberg, Browne, del Rio, Egloff, Garner, Renner, Rieper and Vedral.*

Discussions *with J. Baez, R. Colbeck, B. Fong, J. Vicary and P. Perinotti and many more...*

Funding *from the Swiss National Science Foundation (grant No. 200020-135048), the European Research Council (grant No. 258932) and the National Research Foundation (Singapore).*

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIVERSITY OF OXFORD

NUS
National University of Singapore