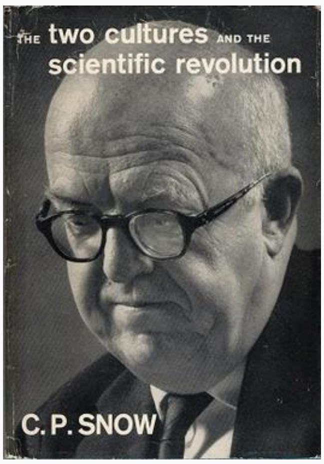
Fluctuation theorems from Bayesian retrodiction

Francesco Buscemi (joint work with Valerio Scarani)

13th Annual Symposium of CQT
National University Singapore (online), 7 Jan 2021

www.quantumquia.com

Have you read a work of Shakespeare's?



"A good many times I have been present at gatherings of people who, by the standards of the traditional culture, are thought highly educated and who have with considerable gusto been expressing their incredulity at the illiteracy of scientists. Once or twice I have been provoked and have asked the company how many of them could describe the Second Law of Thermodynamics. The response was cold: it was also negative. Yet I was asking something which is about the equivalent of: Have you read a work of Shakespeare's?" C.P. Snow (1959)

The "to be or not to be" of thermodynamics

Clausius Inequality

$$\langle \Delta S_{\text{tot}} \rangle \geq 0$$

then throw in, at your discretion, explanations involving: disorder, irreversibility, the arrow of time, life, etc

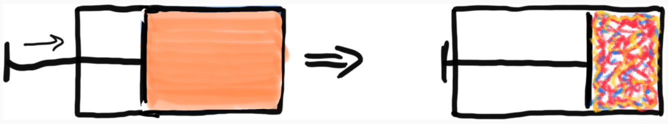
"that is the question"

Indeed!

"No one understands entropy very well."

J. von Neumann (mid/late 1940s)

The Second Law without entropy

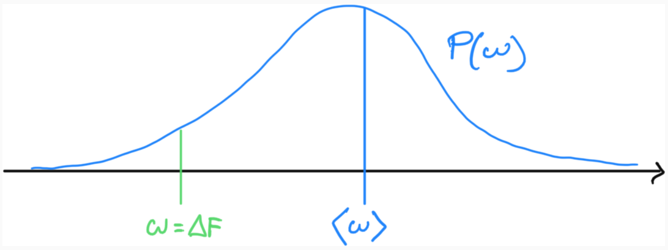


Clausius' inequality (1865):

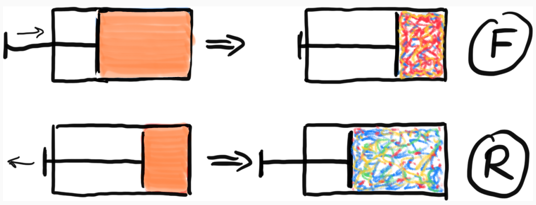
$$\langle W \rangle \geq \Delta F$$

Jarzynski's equality (1997):

$$\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$$

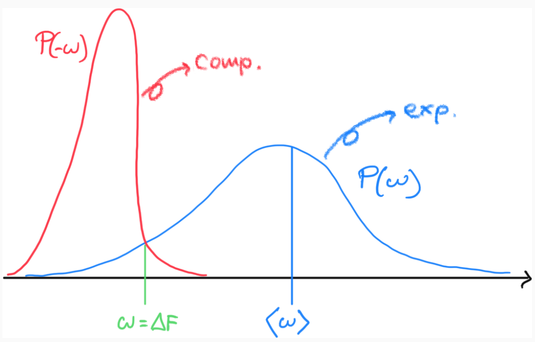


The Second Law and irreversibility



Crooks' fluctation theorem (1999)

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W - \Delta F)}$$



Why is that?

Crooks' theorem relies on two assumptions satisfied at equilibrium:

1. thermal equilibrium: initial distribution is $P(\xi) \propto e^{-\beta\epsilon(\xi)}$
2. microscopic reversibility: molecular processes and their reverses occur at the same rate (viz. probability)

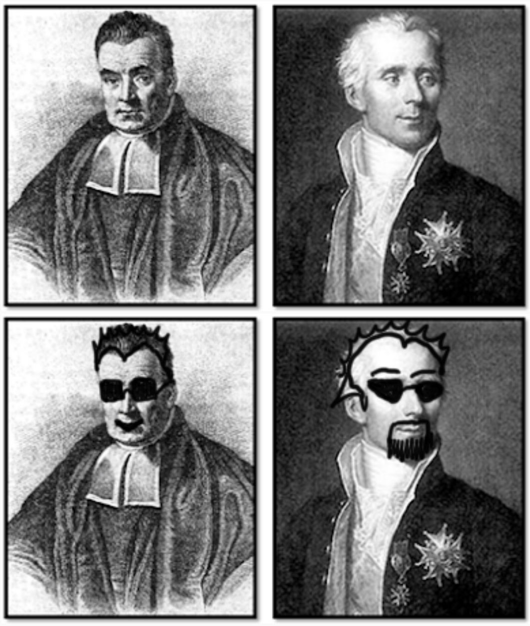
Do we need to know the microscopic details of all the processes involved then?

A hint from Ed Jaynes

"To understand and like thermo we need to see it, not as an example of the n-body equations of motion, but as an example of the logic of scientific inference." E.T. Jaynes (1984)

Reverse process as Bayesian retrodiction

The Bayes-Laplace Rule



Inverse Probability Formula

$$\underbrace{P(H|D)}_{\text{inv. prob.}} \propto \underbrace{P(D|H)}_{\text{likelihood}}\,\underbrace{P(H)}_{\text{prior}}$$

where H is a hypothesis, D is the result of observation (i.e., evidence)

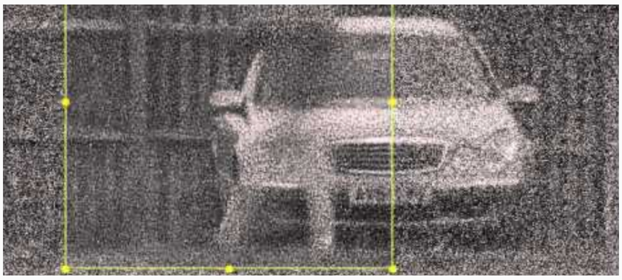
Meaning(s) of the inverse probability

Inverse probability:

- is the main *tool* of Bayesian statistics for problems like:
  - estimation (e.g.: how many red balls are in an urn?)
  - inference and decision (e.g.: is ACME's stock a good investment? should I buy some?)
  - predictive inference (e.g.: weather forecasts)
  - retrodictive inference (e.g.: what kind of stellar event was the supernova of AD 1006?)
- measures the *degree of belief* in one hypothesis among other mutually exclusive ones, given the data

Noisy data and uncertain evidence

- Bayes-Laplace Rule does not tell us what to do with noisy data
- suppose that a noisy observation suggests a probability distribution $Q(D)$ for the data (e.g., the license plate no.)



- how should we update our prior $P(H)$ given *uncertain evidence* $Q(D)$?

Jeffrey's rule of probability kinematics

Belief prediction:

$$P(H) \mapsto P(D) = \sum_H P(D|H)P(H)$$

Belief retrodiction:

$$Q(D) \mapsto Q(H) =?$$

Jeffrey's rule of probability kinematics*

$$Q(D) \mapsto Q(H) = \sum_D \underbrace{P(H|D)}_{\text{inv. prob.}} Q(D)$$

* Jeffrey's rule can be proved from Bayes-Laplace Rule and Pearl's method of virtual evidence

First idea: identify the reverse process with retrodiction

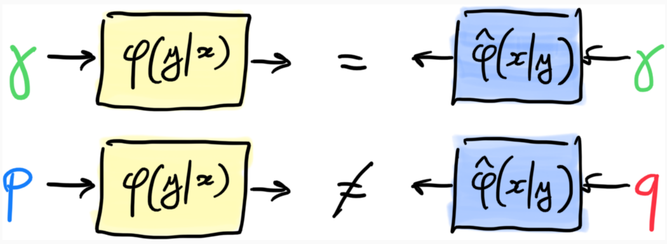
Setup 1/2: construction of the reverse process

- reference process: $\Gamma(x,y) := \varphi(y|x)\gamma(x)$
  - $\rightsquigarrow$ equilibrium condition: $\Gamma(y) = \sum_x \Gamma(x,y) = \gamma(y)$
- Bayesian inversion: $\hat{\varphi}(x|y) := \frac{\Gamma(x,y)}{\Gamma(y)} = \varphi(y|x)\frac{\gamma(x)}{\gamma(y)}$
- at equilibrium, prediction=retrodiction

  $\rightsquigarrow$ i.e., $\varphi(y|x)\gamma(x) = \hat{\varphi}(x|y)\gamma(y)$

Setup 2/2: introducing fluctuations

- no fluctuations at equilibrium: $\varphi(y|x)\gamma(x) = \hat{\varphi}(x|y)\gamma(y)$
- we now change beliefs: $\gamma(x) \mapsto p(x)$ and $\gamma(y) \mapsto q(y)$
- forward process (prediction): $P_F(x, y) := \varphi(y|x)p(x)$
- reverse process (retrodiction): $P_R(x, y) := \hat{\varphi}(x|y)q(y)$
- now, out of equilibrium, $P_F(x, y) \neq P_R(x, y)$



Second idea:
fluctuation relations as measures of "asymmetry" between prediction and retrodiction

Measures of statistical divergence

- ratio: $r(x,y) := \frac{P_F(x,y)}{P_R(x,y)} = \frac{\varphi(y|x)p(x)}{\hat{\varphi}(x|y)q(y)} = \frac{\gamma(y)p(x)}{\gamma(x)q(y)}$
- f-divergences: $D_f(P_F \| P_R) := \sum_{x,y} P_F(x,y) f(r(x,y))$
  - $\rightsquigarrow$ $f(r) = \ln(r) \implies D_f$ is KL-divergence
  - $\rightsquigarrow$ $f(r) = r^\alpha$, $\alpha \neq 0 \implies D_f$ is a Hellinger-Rényi divergence
- here we assume $f : \mathbb{R}^+ \to \mathbb{R}$ smooth and invertible
- thus define $g := f \circ \frac{1}{x} \circ f^{-1}$
  - $\rightsquigarrow$ $f(r) = \ln(r) \implies g(r) = -r$
  - $\rightsquigarrow$ $f(r) = r^\alpha \implies g(r) = \frac{1}{r}$

From f-divergences to f-fluctuation theorems

$D_f(P_F \| P_R) = \sum_{x,y} P_F(x,y) f(r(x,y))$ with $r(x,y) = \frac{\gamma(y)p(x)}{\gamma(x)q(y)}$

- probability density function: $\mu_F^f(u) := \sum_{x,y} \delta[u - f(r(x,y))]\ P_F(x,y)$

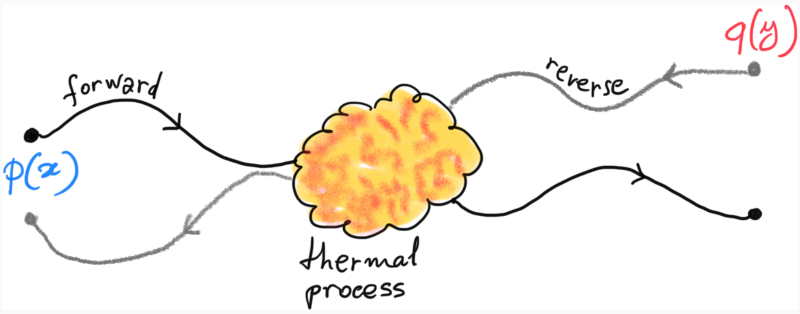
  $\rightsquigarrow$ $\int_{\mathbb{R}} \mu_F^f(u) du = D_f(P_F \| P_R)$
- by consistency: $\mu_R^f(u) := \sum_{x,y} \delta[u - f(\frac{1}{r(x,y)})]\ P_R(x,y)$

f-Fluctuation Theorem

$$\frac{\mu_F^f(u)}{\mu_R^f(g(u))} = \frac{|g'(u)|}{f^{-1}(g(u))} \qquad \implies \qquad \left\langle f^{-1}(g(u)) \right\rangle_F = 1$$

in particular, for $f = \ln$, we have $\frac{\mu_F(u)}{\mu_R(-u)} = e^u$ and $\langle e^{-u} \rangle_F = 1$

Example: recovering work and heat



- $p(x) = e^{\beta(F - E_x)}$
- $q(y) = e^{\beta(F' - E'_y)}$
- $\gamma(x) = e^{\beta(F'' - \epsilon_x)}$

Then, with all due physical assumptions:

$$\begin{aligned}\ln \frac{\gamma(y)p(x)}{\gamma(x)q(y)} &= \beta(F'' - \epsilon_y + F - E_x - F'' + \epsilon_x - F' + E'_y) \\ &= \beta(\underbrace{E'_y - E_x}_{\Delta E} - \underbrace{(\epsilon_y - \epsilon_x)}_{Q} - \underbrace{(F' - F)}_{\Delta F}) \\ &= \beta(W - \Delta F)\end{aligned}$$

Back to "the question"

What is it that *grows* in the Second Law then?

as Jarzynski's equality $\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$ implies Clausius inequality $\langle W \rangle \geq \Delta F$

so the generalized equality $\langle e^{-u} \rangle_F = 1$, obtained for $f = \ln$, implies $D_{\mathsf{KL}}(p\|\gamma) \geq D_{\mathsf{KL}}(\varphi[p]\|\gamma)$

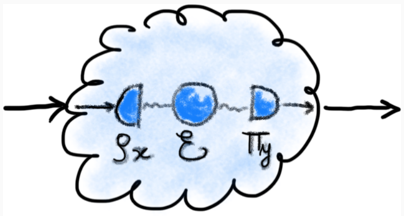
in other words...

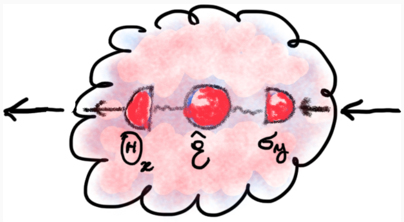
$$\langle \Delta S_{\mathsf{tot}} \rangle \geq 0 \quad \rightsquigarrow \quad D_{\mathsf{KL}}(p^0\|\gamma) \geq D_{\mathsf{KL}}(p^1\|\gamma)$$

i.e., the system gets closer to equilibrium (on average)

Quantum Inside® ready

The case of quantum processes





- assume $\varphi(y|x) = \mathrm{Tr}[\Pi_y\ \mathcal{E}(\rho_x)]$
- according to the formalism of *quantum retrodiction*:
  - $\Gamma := \sum_x \gamma(x)\rho_x$
  - $\sigma_y := \frac{1}{\gamma(y)}\sqrt{\mathcal{E}(\Gamma)}\Pi_y\sqrt{\mathcal{E}(\Gamma)}$
  - $\Theta_x := \gamma(x)\frac{1}{\sqrt{\Gamma}}\rho_x\frac{1}{\sqrt{\Gamma}}$
  - $\hat{\mathcal{E}}(\cdot) := \sqrt{\Gamma}\left\{\mathcal{E}^\dagger\left[\frac{1}{\sqrt{\mathcal{E}(\Gamma)}}(\cdot)\frac{1}{\sqrt{\mathcal{E}(\Gamma)}}\right]\right\}\sqrt{\Gamma}$
- Bayesian inversion carries through easily $\hat{\varphi}(x|y) = \mathrm{Tr}[\Theta_x\ \hat{\mathcal{E}}(\sigma_y)]$

Conclusions

A guide for small talk about the Second Law

When the “entropy always increases” trope does not help, try with these:

- out of equilibrium there is asymmetry between forward and reverse process
- such asymmetry is astronomically ($\approx 10^{23}$) conspicuous in macroscopic situations (the perceived “one-wayness” of time)
- such asymmetry is well described in terms of fluctuation relations (FRs)
- FRs can be seen, not as consequences of complex microscopic balancing mechanisms, but as consequences of Bayes-Laplace Rule
- hence, FRs and the Second Law hold also without physics (e.g.: data-processing theorem in information theory)