Bayesian Retrodiction and the Second Law of Thermodynamics

Francesco Buscemi*

Second Kyoto Workshop on Quantum Inf, Comp, and Foundations

YITP, Kyoto U (online), 13 Sept 2021

*www.quantumquia.com

About these ideas

Two papers:

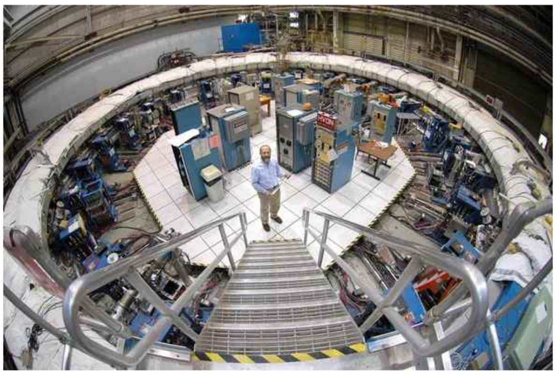
- with V. Scarani. *Fluctuation relations from Bayesian retrodiction*. Phys. Rev. E (2021). arXiv:2009.02849 [quant-ph]
- with C.C. Aw and V. Scarani. *Fluctuation Theorems with Retrodiction rather than Reverse Processes*. AVS Quantum Science (to appear). arXiv:2106.08589 [cond-mat.stat-mech]

New physics!!

Long-Awaited Muon Measurement Boosts Evidence for New Physics

Initial data from the Muon g-2 experiment have excited particle physicists searching for undiscovered subatomic particles and forces

By Daniel Garisto on April 7, 2021 أعرض هذا باللغة العربية



Muon g-2 magnetic storage ring, seen here at Brookhaven National Laboratory in New York State before its 2013 relocation to Fermi National Accelerator Laboratory in Illinois. Credit: Alamy

READ THIS NEXT

PHYSICS
Unexplained Results Intrigue Physicists at World's Largest Particle Collider
March 25, 2021 — Daniel Garisto

PHYSICS
Muons Bring New Physics within Reach
April 13, 2017 — Elizabeth Gibney and Nature

PHYSICS
Muons: The Little-Known Particles Helping to Probe the Impenetrable
May 28, 2018 — Elizabeth Gibney and Nature magazine

New physics?!?

MailOnline Science & Tech

Home | News | U.S. | Sport | TV&Showbiz | Australia | Femail | Health | Science | Money | Food | Travel | Best Buys | Discounts

Latest Headlines | NASA | Apple | Twitter Login

A step towards LIMITLESS energy: Loophole found in a fundamental law of physics may lead to infinite power

- The finding may mean it's possible to create perpetual motion machines
- These machines can spin for eternity without losing any energy
- The four laws of thermodynamics set the physical rules for our universe
- Researchers found a way to bypass the second law of the four
- They have since projected a quantum system in which energy can be recycled

By HARRY PETTIT FOR MAILONLINE
PUBLISHED: 14:00 BST, 3 November 2016 | UPDATED: 17:03 BST, 3 November 2016

Share 2.4k shares 282 View comments

Einstein once boldly claimed that the Laws of Thermodynamics were the only physical theory of the universe that will 'never be overthrown'.

That all changed late last month, when scientists from the **Argonne National Laboratory** at the University of Chicago found a loophole in the system - one that allows them to break the second law of thermodynamics.

The Second Law is "special"

"The law that entropy always increases holds, I think, the supreme position among the laws of Nature. [...] If your theory is found to be against the Second Law of Thermodynamics I can give you no hope; there is nothing for it to collapse in deepest humiliation."

A.S. Eddington

"[...] the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown." A. Einstein

The statement

The Second Axiom of Thermodynamics

A *perpetuum mobile* of the second kind* is impossible. In formula,

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

* A machine that extracts work from a single heat bath.

Why does the above "feel" so special among physical laws?

Is entropy the key?

Many "explanations" of the Second Law actually amount to explanations of the meaning of entropy (e.g., counting arguments).

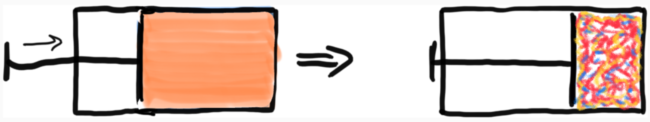
Problem is...

" No one understands entropy very well..."

von Neumann (apocryphal)

" ...and that's only half of the story, anyway." Anon

The Second Law without entropy

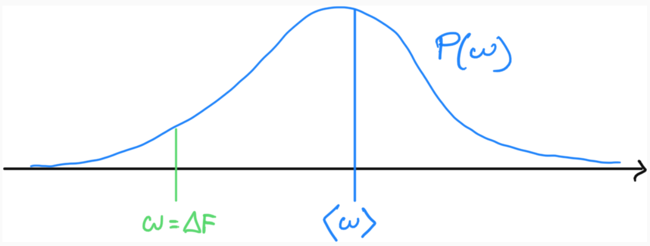


Clausius' inequality (1865):

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

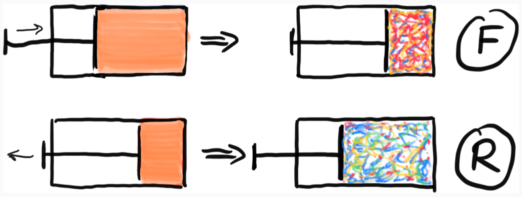
Jarzynski's equality (1997):

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



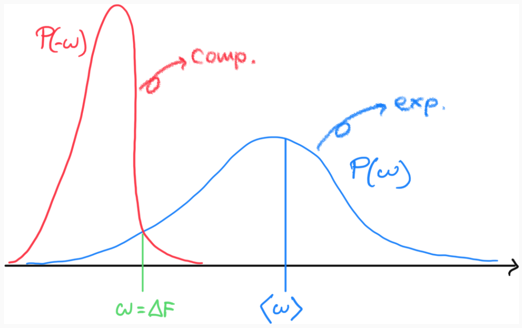
Jarzynski $\Longrightarrow$ Clausius

Enter irreversibility



Crooks' fluctation theorem (1999)

$$\frac{\mathcal{P}_F(W)}{\mathcal{P}_R(-W)} = e^{\beta(W-\Delta F)}$$



Crooks $\Longrightarrow$ Jarzynski $\Longrightarrow$ Clausius

Usual explanation

Crooks' theorem, and hence Jarzynski's relation, and hence the Second Law, all rely on two assumptions satisfied at equilibrium:

1. thermal distribution: microstate probability is $\mathcal{P}(\xi) \propto e^{-\beta\epsilon(\xi)}$
2. microscopic reversibility (cf. *detailed balance*): molecular processes and their reverses occur at the same rate

So, is the Second Law special because of some kind of "special" microscopic balancing mechanism?

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****."*

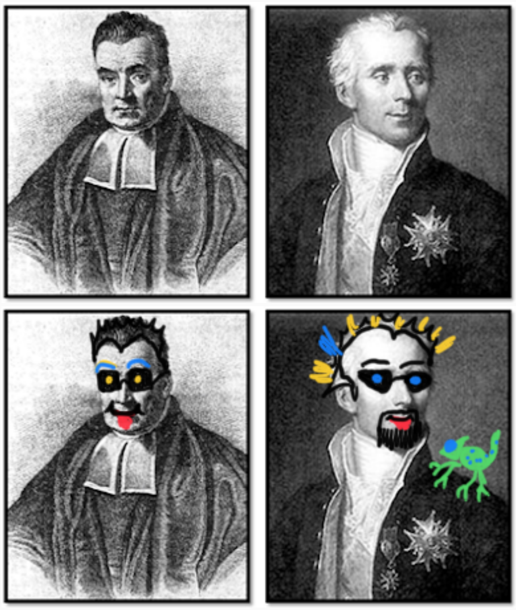
E.T. Jaynes (1984)

A hint from Satosi Watanabe

"The phenomenological oneway-ness of temporal developments in physics is due to **irretrodictability**, *and not due to irreversibility."* S. Watanabe (1965)

Reverse process as Bayesian retrodiction

The Bayes-Laplace Rule



Inverse Probability Formula

$$\underbrace{\mathcal{P}(H|D)}_{\text{inv. prob.}} \propto \underbrace{\mathcal{P}(D|H)}_{\text{likelihood/model}} \; \underbrace{\mathcal{P}(H)}_{\text{prior}}$$

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

postmodern Bayesianism!

Meanings of inverse probability

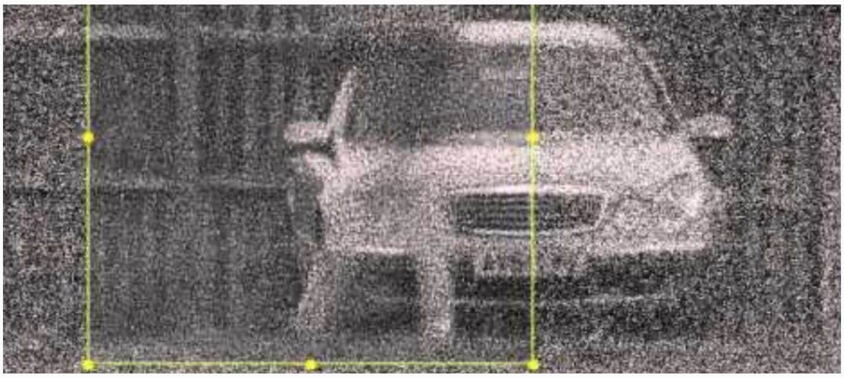
It is the main *tool* of Bayesian statistics for problems like:

- **estimation** (e.g.: how many red balls are in an urn?)
- **decision** (e.g.: is ACME's stock a good investment? should I buy some? how much?)
- **inference and learning**:
  - predictive inference (e.g.: weather forecasts)
  - retrodictive inference (e.g.: what kind of stellar event possibly caused the Crab Nebula?)

Inference with noisy data or uncertain evidence

BUT! Bayes-Laplace Rule does not tell us how to update the prior in the face of *uncertain* data...

- suppose that a noisy observation suggests a probability distribution $\mathcal{Q}(D)$ for the data (e.g., the license plate no.)



- how should we update our prior $\mathcal{P}(H)$ given *uncertain evidence* in the from $\mathcal{Q}(D)$?

Jeffrey's rule of probability kinematics

Vanilla Bayes:

$$\mathcal{P}(H|D) = \mathcal{P}(D|H)\mathcal{P}(H)/\mathcal{P}(D)$$

Extended Bayes:

$$\mathcal{P}(H|\mathcal{Q}(D)) = ?$$

Jeffrey's conditioning* (1965)

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

$$= \sum_D \frac{\mathcal{P}(D|H)\mathcal{P}(H)}{\mathcal{P}(D)} \mathcal{Q}(D)$$

* Jeffrey's rule was introduced *ad hoc*, but it can be proved from Bayes-Laplace Rule and Pearl's method of virtual evidence (1988)

Jeffrey's rule "promotes" Bayes inverse probability to a fully fledged channel

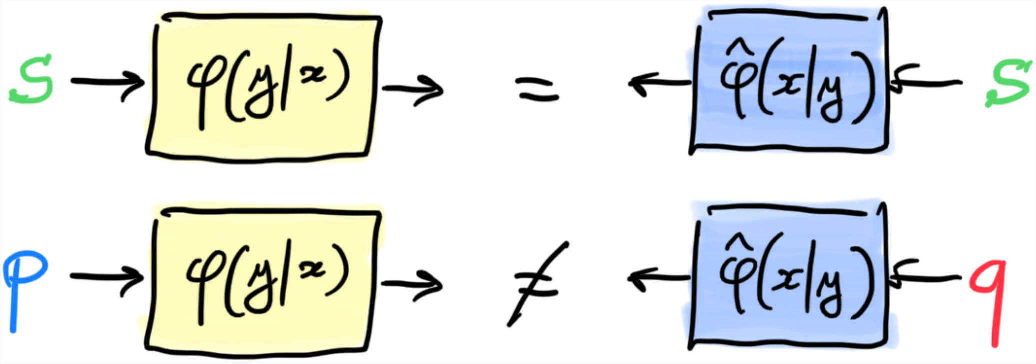
Construction of the reverse process as retrodiction

- **physical setup**:
  - a stochastic transition rule: $\varphi(y|x)$
  - a steady (viz. invariant) state: $\sum_x \varphi(y|x)s(x) = s(y)$
- **Bayesian inversion at the steady state**:

$$s(y)\hat{\varphi}(x|y) := s(x)\varphi(y|x) \quad \Longleftrightarrow \quad \frac{\varphi(y|x)}{\hat{\varphi}(x|y)} = \frac{s(y)}{s(x)}$$

- **two priors**:
  - predictor's prior: $p(x)$
  - retrodictor's prior $q(y)$
- **two processes**:
  - forward process (prediction): $\mathcal{P}_F(x,y) = \varphi(y|x)p(x)$
  - reverse process (retrodiction): $\mathcal{P}_R(x,y) = \hat{\varphi}(x|y)q(y)$

A picture



- at the steady state: prediction $=$ retrodiction
- otherwise: asymmetry (irreversibility, *irretrodictability*)

Fluctuation relations as measures of irretrodictability

Quantifying irretrodictability

- relative entropy:

$$D(\mathcal{P}_F \| \mathcal{P}_R) := \left\langle -\ln \frac{\mathcal{P}_R(x,y)}{\mathcal{P}_F(x,y)} \right\rangle_F =: \langle -\ln r(x,y) \rangle_F$$

$\rightsquigarrow$ more generally, one can use $D_f(\mathcal{P}_R \| \mathcal{P}_F) := \langle f(r(x,y)) \rangle_F$

f-Fluctuation Theorem

$$\mu_R(\omega) = f^{-1}(\omega)\mu_F(\omega) \quad \Longrightarrow \quad \langle f^{-1}(\omega) \rangle_F = 1$$

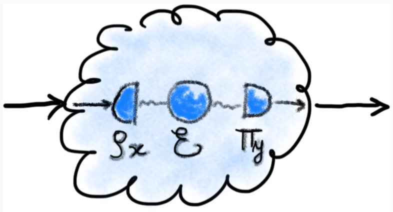
$\rightsquigarrow$ for $f(u) = -\ln u$, we have $f^{-1}(v) = e^{-v}$, that is

$$\frac{\mu_F(\omega)}{\mu_R(\omega)} = e^{\omega} \quad \Longrightarrow \quad \left\langle e^{-\omega} \right\rangle_F = 1$$

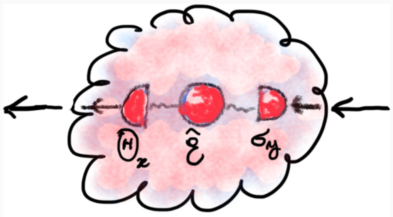
Example: nonequilibrium steady states

- stochastic process $\varphi(y|x)$ with non-thermal steady state $s(x)$
- thermal equilibrium priors: $p(x) = q(x) \propto e^{-\beta\epsilon_x}$
- fluctuation variable:
  $\omega = \ln \frac{\mathcal{P}_F(x,y)}{\mathcal{P}_R(x,y)} = \ln \frac{p(x)}{q(y)} \frac{s(y)}{s(x)} = \beta(\epsilon_y - \epsilon_x) + (\ln s(y) - \ln s(x))$
- *nonequilibrium potential*: $V(x) := -\frac{1}{\beta} \ln s(x)$ (e.g., Manzano&al 2015)
- nonequilibrium potentials (usually introduced *ad hoc*) are understood here as remnants of Bayesian inversion
- $\Longrightarrow \left\langle e^{\beta(\Delta E - \Delta V)} \right\rangle_F = 1 \Longrightarrow D(p\|s) - D(\varphi[p]\|s) \geq 0$

Example: Quantum Inside©



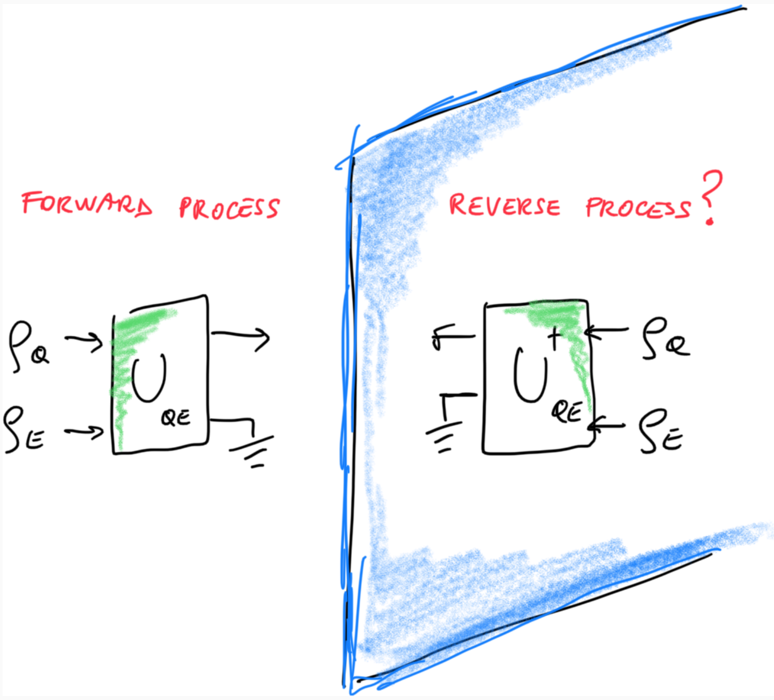
- assume $\varphi(y|x) = \mathrm{Tr}[\Pi_y \, \mathcal{E}(\rho_x)]$
- let $s(x)$ be invariant distribution
- perform *quantum retrodiction*:
    - $\Sigma := \sum_x s(x)\rho_x$
    - $\hat{\rho}_y := \frac{1}{s(y)}\sqrt{\mathcal{E}(\Sigma)}\Pi_y\sqrt{\mathcal{E}(\Sigma)}$
    - $\hat{\Pi}_x := s(x)\frac{1}{\sqrt{\Sigma}}\rho_x\frac{1}{\sqrt{\Sigma}}$
    - $\hat{\mathcal{E}}(\cdot) := \sqrt{\Sigma}\left\{\mathcal{E}^\dagger\left[\frac{1}{\sqrt{\mathcal{E}(\Sigma)}}(\cdot)\frac{1}{\sqrt{\mathcal{E}(\Sigma)}}\right]\right\}\sqrt{\Sigma}$
- Bayes–Jeffrey inversion works seamlessly



$$\hat{\varphi}(x|y) = \mathrm{Tr}[\hat{\Pi}_x \, \hat{\mathcal{E}}(\hat{\rho}_y)]$$

The origin of irretrodictability

The problem with the notion of “reversal”



What sort of transformation is it? Is it always well-defined? How is it implemented?

“Physical transformation” or “belief propagation”?

Not “objective”. In stat-mech, the construction of the reverse process depends on a *choice* of system-bath interaction and reference prior.

Not “constructive”. Even if a physical realization (e.g., a circuit implementation) of the forward process is available, that does not mean that its reverse is also physically available.

$\Longrightarrow$ the reverse process does not depend only on the forward process, but also on the agent’s belief!

$\Longrightarrow$ prediction and retrodiction are fundamentally different: origin of a logical/inferential arrow.

Special case: Hamiltonian processes

The following are equivalent (both in classical and quantum theory):

- a given process is Hamiltonian
- its reverse does not depend on the choice of prior
- it is *bilaterally deterministic*

Interpretation

The reverse process is agent-independent if and only if the process is Hamiltonian.

$\Longrightarrow$ a reversal **always exists**; however, it is agent-independent for, and only for, Hamiltonian processes

Conclusions

Final messages

Conceptual insights:

1. one-way-ness: not irreversibility, but irretrodictability
2. entropy increase: not "time arrow", but "inferential arrow"
3. reversal: not physical transformation, but Bayesian inversion
4. $\Longrightarrow$ the Second Law is special among physical laws because it is not so much a law of physics, as it is a law of logic

Applications:

1. fluct. relations without "ad hockeries" e.g. non-eq. potentials
2. fluct. relations and Second Law beyond thermo and physics

thank you