Some Misconceptions about Entropy

- Introduction — the ground rules.
- Gibbs versus Boltmann entropies.
- That awful $H$-theorem.
- The Second Law of Thermodynamics.
- Tsallis and other heresies.
- Open the quantum box.
- Time asymmetry and the approach to equilibrium.
- Conclusions.
Some Misconceptions about Entropy

**THERMODYNAMICS AND STATISTICAL MECHANICS**

Enthropy

Clausius (1850) defined entropy as a state variable for systems in equilibrium.

\[
\Delta S_E = \int \frac{dQ}{T} \quad \text{reversible}
\]

**Experimental Enthropy**

Function of macroscopic variables.

Classical Thermodynamics is the result. Conceptually clear. No problems.

The problem arose when trying to give microscopic, statistical interpretation.

Statistical Thermodynamics.

Boltzmann's (1866) Kinetic Theory considered \( N \) particles distributed in 6-D phase space, and how collisions lead to equilibrium distributions. He defined \( H \) function as

\[
H = \int d\mathbf{p} \mathit{e} \ln \mathit{e} \quad \text{where} \quad \mathit{e}(\mathbf{r}, \mathbf{p}, t) \text{ is distribution of particles}
\]
**STATISTICAL MECHANICS — GIBBS VS. BOLTZMANN**

Gibbs' Statistical Mechanics. (1878)

Focussed attention on the total system of N particles in GND phase space.

To introduce statistical notions he used the artifice of the ensemble, a large number of (mental) copies of the system.

Gave us the **GIBBS’ ALGORITHM** to setup ensemble.

Maximise $S_G = -k \int dx \ p_N \log p_N$

under available macroscopic constraints eg: $\langle F \rangle = \int F p_N$

where $p_N$ is distribution function for N-particle system.

This method is successful to this day. [Con: Trellope]

We teach a strange mixture:

1. Use Gibbs’ Algorithm for calculation.
2. Try to justify it using the language of Boltzmann.

There are quite a few misconceptions still current and stem from a misunderstanding of the role of probability in statistical mechanics.
Inference — A Bayesian Perspective

Inference — The Ground Rules. A Bayesian viewpoint.

Three separate stages in inductive inference.

1) **Calculation** — Bayes' Theorem

\[ \text{pr}(A, B) = \text{pr}(A) \cdot \text{pr}(B | A) \]

↑
A and B

↑
A given B

Any consistent assignment of real numbers to any propositions A and B must obey the rules of (Bayesian) probability (Cox 1946).

2) **Assignment**:

Probabilities (above) can only be assigned consistently by Maximising the ENTROPY:

\[ S = - \sum \text{pr}(i) \cdot \log \text{pr}(i)/m(i) \]

where \( \text{pr}(i) = \text{prob of } i \text{th possible state.} \)
\( m(i) = \text{measure (no. of microstates).} \)

Under available constraints on \( \text{pr}(i) \), eg \( \langle r \rangle = \sum_i r_i \text{pr}(i) \)
LAPLACE, GIBBS, JAYNES, SHORE & JOHNSON, SKILLING.

3) **Finding the right hypothesis space**.

(Enumerate the possible states of the system (DYNAMICS))

This is our task as physicists.
Most physicists live entirely at level (3): studying dynamics, building models for systems. — "real physics'.

**MODELS FOR REALITY = ONTOLOGY**

Statistical Mechanics (and image processing!) works almost entirely at levels (2) and (1): KNOWLEDGE ABOUT REALITY = EPISTEMOLOGY

In particular, Gibbs’ “ENSEMBLE” defines the probability that the N-particle system is in a particular state. We are making consistent inferences about the state of the system, given incomplete information (macroscopic variables).

We are not assuming that the system actually explores all the states accessible within the constraints, or even that it changes state at all. [No ergodic assumption].

If the constraints available lead to good predictions, Good. If not: What are the other influences?
Some Misconceptions about Entropy

GIBBS VS. BOLTZMANN ENTROPIES

\[ \frac{N}{\text{classical, interacting particles in box volume } V} \]
\[ \text{Hamiltonian } \sum \frac{p_i^2}{2m} + U(x_1, \ldots, x_N) \]

GIBBS’ \[ S_G = -k \int d\nu \ p_N \log p_N \]

BOLTZMANN’s \[ S_B = -kN \int d\nu \ p_i \log p_i \]
(suitably re-interpreted)

\[ p_i = \text{1-particle marginal distribution } \int d\nu \ p_N \]

Which is meaningful? Gibbs’, of course.

You haven’t got \( N \) systems of 1 particle, you have 1 system of \( N \) particles.

Furthermore, I can prove that, if the system is in the state of maximum \( S_G \) (canonical ensemble) and if the state variables are altered over a locus of equilibrium states (reversible paths) then:

\[ \Delta S_G = \int \frac{dQ}{T} \]

but:

\[ \Delta S_B = \int d\langle k \rangle + p \cdot dV \quad \text{(ignores internal energy)} \]
\[ \quad \text{(and interparticle forces)} \]
\[ \quad k = \text{kinetic energy} \]
Hence: The Gibbs' entropy when maximised

\[ \Delta S_G = \int_1^2 \frac{dQ}{T} = \Delta S_E \]

and so can be identified with the experimental entropy defined by Clausius.

More generally, because \( S_G \) is defined for all distributions, not just the canonical ensemble,

\[ S_G \leq S_E \]

with equality iff \( p_\text{N} \) is canonical.

The Boltzmann Entropy \( S_B \) ignores the internal energy and the effect of inter-particle forces on the pressure. And is not equal to the experimental \( S_E \) except for a perfect gas [when it equals \( S_G \) as well].

**MORAL:** \( S_G \) is the correct theoretical concept. \( S_B \) is NOT USEFUL.
THAT AWFUL H-THEOREM IN ALL THE BOOKS

Shows that $S_B$ always increases:

$$\frac{dS_B}{dt} = -k \sum \log \frac{p_i}{\bar{p}_i} \frac{dp_i}{dt}$$

Now imagine the 1-particle systems making quantum jumps between the 1-particle microstates. Use the detailed balance equation to show:

$$\frac{dS_B}{dt} = k \sum_{qp} \nu_{qp} \left( \log p_{\beta q} - \log p_{\alpha q} \right) \left( p_{\beta q} - p_{\alpha q} \right) \geq 0$$

UGH!! Several things that are wrong.

(1) Approximate quantum mechanics. Not necessarily valid for large perturbations.

(2) Bad quantum mechanics. The N-particle system has N-particle states. It will sit in one of them, and not make transitions at all.

(3) Even if you could prove it, it wouldn’t be useful unless $\int \frac{dS_E}{dt} dt = \Delta S_E$ [it isn’t].

(4) There are whole classes of COUNTER-EXAMPLES: eg free expansion of O2 at 45 atm and 160 K.
THE SECOND LAW OF THERMODYNAMICS

The psychological need for an H-Theorem is of course related to a misconception about the 2nd Law of Thermodynamics.

For isolated system: \( \Delta S_E > 0 \)

equality iff reversible.

The misconception is that theoretical entropy has to do so too. In fact because:

EITHER (1) of Liouville’s Theorem. (classical)
OR (2) System in N-particle eigenstates. (quantum)

The Gibbs’ Entropy \( S_G \) is a constant of motion.

Remarkably, this is just what you need to PROVE the second law of Thermodynamics.
Some Misconceptions about Entropy

PROOF OF THE SECOND LAW OF THERMODYNAMICS

For special case. God of N particles.

Box with partition at $t=0$ confined to one side.

(1) Suppose system has canonical probability distribution $p_i \propto e^{-\beta E_i}$

\[ \text{in equilibrium with } S_G = S_E \]

(2) Open partition and wait until state variable stop changing. So in “equilibrium” so $S_E$ can be defined.

(3) But all motions Hamiltonian, so $S_G' = S_G$ has not changed.

(4) But probability distribution of N-particles NO LONGER CANONICAL, because of (very subtle!) correlations that express the fact it was once on one side of the partition. So:

\[ S_G' \leq S_E' \]

So, \[ S_E \leq S_E' \] Q.E.D.

THE SECOND LAW OF THERMODYNAMICS IS A LAW ABOUT EXPERIMENTAL QUANTITIES.
Some Misconceptions about Entropy

**THEORETICAL SECOND LAW**

Planck (1906)

Boltzmann was right after all....

\[ S = k \log W \]

⇒ Gibbs' Entropy Sc.!

\(W\) = number of microstates compatible with macroscopic state.

STATE I.

\(T_1, p_1,\) etc.

STATE II.

\(T_2, p_2,\) etc.

Fundamental requirement for reproducible process!

Phase Volume compatible with Final state cannot be less than phase volume of the initial state (which describes our ability to reproduce the initial state).

This is also the 2nd Law.

Equality of phase volumes => Reversible
ENTROPY INCREASE AND THE ARROW OF TIME

- Boltzmann’s constant \((k_B = 1.38 \times 10^{-23} \text{ J K}^{-1})\) is rather small, and the increase of phase-space volume is usually very large.

- (Taken from a Cambridge Part IB Physics Example Sheet.) Suppose that an infrared photon of energy 1 eV is absorbed by a dust grain at 300 K. The entropy increases by \(5.34 \times 10^{-22} \text{ J K}^{-1}\).

- This increases the available phase-space volume of the Universe by a factor of \(\exp(38.7) = 6.3 \times 10^{16}\).

- Scaled up to the size of this lecture theatre, we get a staggering \(\exp(10^{21})\) increase in volume PER SECOND.

- Macroscopic irreversibility is therefore not surprising in the least...
There has been a lot of papers published concerning the Tsallis generalised entropy.

These are one-parameter family of functions based on the \( q \)-derivative, and include the correct entropy as a special case \( q = 1 \).

The “entropic” functions for other values of the parameter \( q \) are non-extensive (i.e. do not satisfy the Kangaroo axiom).

The maximised “entropies” for \( q \neq 1 \) are not equal to the experimental entropy defined by Clausius.

The industry of \( q \)-generalisations has taken hold in many different fields. It is mathematically self-consistent, and may be great fun, but its track record for concrete achievements is (in my opinion) still zero.

That said, Constantino Tsallis has published a very impressive list of applications throughout physics and astrophysics.

But I don’t believe him...
Quantum particle in ground state of 1-dimensional box $\frac{-1}{2}a < x < \frac{1}{2}a$. At $t = 0$ the box is opened to $-a < x < a$.

We can expand the old ground state in terms of the new states. It is still in a pure state, though not an energy eigenstate, and the entropy is still zero.

Evolution of box wavefunction after a long time. The probability either has a central hump, or two symmetrical ones, as the wavefunction displays the interference of the 2 lowest frequency modes.
• A movie shows that the particle oscillates around enjoying its new-found freedom. It doesn’t remotely settle down. Some authors average the phases and thereby get an entropy increase of 0.683714.

• The time average (500 samples) of the probability (red line) is almost indistinguishable from the phase average.
• If the box is opened one-sidedly the oscillations are more violent. The averaged distribution has two humps and the phase-averaged entropy increase is 1.03500.

• But this is NOT thermodynamics.
The canonical distribution with the same average energy has an entropy of 1.594.

If we impose symmetry on the wavefunction the entropy increase is 1.0414 and there is a hump in the middle.
Where does time asymmetry enter physics?

**Answer:** It does not enter via the dynamics. It enters because our information about the system is not time-symmetric. This is then properly reflected in our inferences about its likely behaviour.

**Gibbs’ Algorithm** is a wonderful way of setting up an equilibrium ensemble. How must it be modified to cope with non-equilibria?

**Answer:** The formalism is fine just as it is. Just give the formalism some time-dependent information and it will predict how the system is likely to evolve and approach equilibrium.

\[
\begin{align*}
\langle e^2 \rangle \quad \text{given} & \quad \Rightarrow \quad e \propto e^{-\frac{1}{2}A} \\
\langle e^2(t) \rangle \quad \text{given} & \quad \Rightarrow \quad e \propto e^{-\int s(x,t)A(x,t)dt}
\end{align*}
\]

For more details see (1) Jaynes: Where do we stand on MaxEnt? (in collected papers).
Some Misconceptions about Entropy

TIME ASYMMETRY AND NON-EQUILIBRIA

A case study - Brownian Motion.

Particle mass moves in 1-Dimension
subject to random collisions from molecules at temperature T.

Adopt a slightly different approach to cope
with outside influences. [Dynamics enters via constraint]
[Not Hamiltonian]

State of knowledge described (classically)
by $p_0 \{ \{ x(t) \} \}$ [a big space]

For simplicity: restrict attention to position at a set
of times on grid: $x_i = x(t_i)$ $t_i = i \Delta$

Define average velocity: $v_i = (x_{i+1} - x_i) / \Delta$
acceleration: $a_i = (x_{i+1} - 2x_i + x_{i-1}) / \Delta^2$

Set up "equilibrium" ensemble by maximising ENTROPY

$S(p(x)) = - \int dx \ p(x) \log p(x)$

[Uniform measure in $x$]

under suitable constraints.
Some Misconceptions about Entropy

Brownian Motion

Suitable constraints: for equilibrium.

1. \[ \langle \sigma_i^2 \rangle = \frac{kT}{m} \quad \text{for all } i \quad (t_i) \]

or even the weaker constraint:

\[ \frac{1}{N} \sum_i \langle \sigma_i^2 \rangle = \frac{kT}{m} \]

2. (Because the colliding particles can only provide a certain average impulse \( P \) in time \( \tau \))

\[ \frac{1}{N} \sum_i \langle a_i^2 \rangle = \left( \frac{P}{m} \right)^2 \]

NB: The specification of \( P \) as function of \( \tau \) is dynamics. We are assuming that one \( P, \tau \) pair is sufficient to describe the reproducible features of this system. [This may or may not be adequate.]

3. \[ \frac{1}{N} \sum_i \langle x_i^2 \rangle = L^2 \]

[to keep the particle some where in Cambridge.]

Oxford
**EQUILIBRIUM ENSEMBLE FOR BROWNIAN MOTION**

\[
pr(x) = Z(\alpha, \beta, \gamma) \exp \left( -\frac{1}{2} x^T [\alpha + \beta vv' + \gamma AA^T] x \right)
\]

where \( V \) and \( A \) are velocity/accel. matrices, \( \alpha, \beta, \gamma \) are Lagrange multipliers.

Partition Function:

\[
Z(\alpha, \beta, \gamma) = \int d^x \exp \left( -\frac{1}{2} x^T [\alpha + \beta vv' + \gamma AA^T] x \right)
\]

\( \alpha, \beta, \gamma \) determined by usual methods. \([z]-\text{transform}\)

Note:

\[
pr(x) \propto \exp \left( -\frac{1}{2} x^T \hat{R}^{-1} x \right)
\]

which is: probability distribution for:

zero mean, multivariate, correlated Gaussian time-series

with covariance matrix \( \langle \Delta x \Delta x' \rangle = R \)

The average value \( \langle x_i \rangle = 0 \) for all \( t \)

because the information given is time-independent.
Some Misconceptions about Entropy

**Brownian Motion and Time Asymmetry**

Just provide more information: \((\text{Data} \equiv D)\)

\[
\begin{align*}
 t &= t_1, \quad x = x_1 \pm \delta_1, \\
 t &= t_2, \quad x = x_2 \pm \delta_2
\end{align*}
\]

**Bayes' Theorem**

\[
\Pr(x|D) \Pr(D) = \Pr(x) \Pr(D|x)
\]

- **Constant**
- **Prior equilibrium ensemble**
- **Likelihood** \(\exp -\frac{1}{2}(x - x_i)^2\)

The resulting posterior ensemble shows evolution of \(\langle x(t) \rangle\) and \(\langle \Delta x(t) \rangle\) for both forward and backward in time.

1. **One data point** - Time symmetric, no net flux
2. **Two data points** (light particle) - Asymmetric, net flux appears.
3. **Velocity information** (massive particle) - Slows to zero velocity approaching equilibrium.

There are general lessons about non-equilibrium stat. mech. in this simple example.
Equilibrium ensemble. Quantity $f$

We calculate $\langle f \rangle$ and $\Delta f^2 = \langle (f - \langle f \rangle)^2 \rangle$

$\Delta f^2$ is our uncertainty about $f$.

But is it fluctuating?

Maybe! $\overline{f} = \frac{1}{T} \int_0^T f(t) \, dt$ $\Delta f^2 = \frac{1}{T} \int_0^T (f(t) - \overline{f})^2 \, dt$

This is related to 'ergodic' concepts.

Certainly: $\langle \overline{f} \rangle = \langle f \rangle$

and $\langle (\Delta f)^2 \rangle = (\Delta f)^2 + (\Delta \overline{f})^2$

$\Delta \overline{f}$ is not necessary. $= 0$

To find out if these estimates are reliable we have to go to higher-order correlations....

Non-trivial justification required!

Gibbs’ algorithm gives uncertainty. Not fluctuations.
CONCLUSIONS

(1) Statistical Physics is the science of making inferences about a physical system.

(2) Sort out the **DYNAMICS** from the **INFERENCES**.
- Microstates
- Hamiltonians
- **MAXENT** / **BAYES**
- **ONTOLOGY** vs. **EPISTEMOLOGY**

(3) You have 1 system of N-particles, not N systems of 1 particle.

(4) Distinguish between Gibbs’ and Boltzmann Entropies. — The Gibbs’ form is correct.

(5) There’s no such thing as an H-Theorem.

(6) The 2nd law is a relation between experimental quantities, or phase volumes.

(7) Time asymmetry enters physics via **INFERENCES** not **DYNAMICS**.

(8) The Gibbs’ Algorithm will run and run.
Some Misconceptions about Entropy

APPENDIX — PROOF OF $S_G = S_E$

$$S_B - S_G = -k \int dx \ p_N \log \frac{p_i(1) \ldots p_i(N)}{p_N} \geq 0$$

$\Rightarrow$ THEOREM. $S_G \leq S_B$ with equality iff $p_N$ independent.

Canonical ensemble $p_N \propto \exp^{-\beta H}$

$$p_N = \left(\frac{\hbar}{2\pi m}\right)^{3N/2} Q^{-1} \exp^{-\beta U} - \beta \sum_i \frac{p_i^2}{2m}$$

$$Q(\beta, V) = \int_V \exp^{-\beta U} d^3x_1 \ldots d^3x_N = V \int d^3x, \exp^{-\beta U}$$

1-particle $p_i = \left(\frac{\hbar}{2\pi m}\right)^{3/2} V^{-1} \exp^{-\beta p_i^2/2m}$

$S_B - S_G = -k \left[ \log Q - N \log V + \beta \langle U \rangle \right]$  

But $\langle U \rangle = -\frac{\partial \log Q}{\partial \beta}$  

$\beta \langle p \rangle = \frac{\partial \log Q}{\partial V}$  

[Standard Partition function manipulation]

$. \ \ \ \text{on small change:}$

$$d(S_B - S_G) = -k \left[ \beta d\langle U \rangle + \beta (\langle p \rangle - p_0) dV \right]$$

where $p_0 = \frac{N k T}{V}$
Appendix — Proof of $S_G = S_E$ II

Proof of meaning of $\Delta S$ (II)

$$\Delta (S_B - S_G) = - \int_1^2 \frac{d\langle u \rangle + \langle p \rangle - p_0}{T} \, dV$$

But $S_B = -k \left[ \frac{3}{2} Nk \log(2\pi m kT) + Nk \log(V) + \frac{3}{2} Nk \right]$.

$$\left( \frac{\partial S_B}{\partial T} \right)_V = \frac{3Nk dT}{T} = \frac{d\langle k \rangle}{T} \quad (\text{kinetic energy}) \quad \langle p^2 \rangle = 3nkT$$

$$\left( \frac{\partial S_B}{\partial V} \right)_T = \frac{Nk dV}{V} = \frac{p_0 dV}{T}$$

$$\Rightarrow \Delta S_B = \int_1^2 \frac{d\langle k \rangle + p_0}{T} \, dV$$

$$\Delta S_G = \int_1^2 \frac{d(\langle k \rangle + \langle u \rangle) + \langle p \rangle}{T} \, dV$$

$$= \int_1^2 \frac{dQ}{T}$$
Some Misconceptions about Entropy

BROWNIAN MOTION

Brownian Motion

\[ \alpha = 0 \]
\[ \gamma / \beta \approx 1 \]
'Light' particle.

\[ \alpha = 0 \]
\[ \gamma \rightarrow 0 \]
'Light' particle.

\[ \alpha \rightarrow 0 \]
\[ \gamma / \beta \approx 20 \]
'Heavy' particle.