# Key Concepts and Problems in the Philosophy of Thermodynamics

Lorenzo Lorenzetti
USI

14 April 2026
USI, Lugano

# What is thermodynamics?

- Study of heat, work, and macroscopic change
- Born in the context of steam engines
- Now applies to macroscopic systems generally

$$T,\ P,\ V,\ U,\ S$$

temperature pressure volume internal energy entropy

# A phenomenological theory

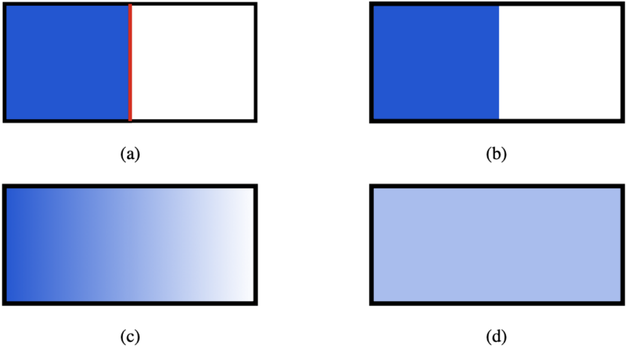
- Classical thermodynamics is a **phenomenological** theory
- It describes systems through macroscopic variables
- These are well-defined at or near **equilibrium**

$$\frac{dX}{dt} \approx 0$$

At equilibrium, macroscopic properties remain approximately stable.

# Minus first law: approach to equilibrium

- Isolated systems typically relax toward equilibrium
- Equilibrium states are macroscopically stable: thermodynamic parameters determined by internal energy U and by external parameters (e.g. volume V)



(a) (b) (c) (d)

# **Zeroth law: thermal equilibrium**

- Thermal equilibrium is transitive
- This makes temperature a well-defined quantity

$$A \sim B, \qquad B \sim C \qquad \Rightarrow \qquad A \sim C$$

$$T_A = T_B \qquad \text{in thermal equilibrium}$$

This is what makes thermometry possible.

# **First law**: conservation of energy

- Internal energy changes through heat and work

$$\Delta U = Q + W$$

$$dU = \delta Q + \delta W$$

Energy is neither created nor destroyed.

# Work and equations of state

- For a simple compressible system:

$$\delta W = -P\,dV$$

- More generally:

$$\delta W = -\sum_i P_i\,dV_i$$

- Concrete applications require equations of state

$$PV = nRT$$

# Second law: traditional statements

Eddington: *“if your theory is found to be against the Second Law of Thermodynamics I can give you no hope”*

**Kelvin**

- No process has as its sole effect the complete conversion of heat from a single source into work

**Clausius**

- No process has as its sole effect the transfer of heat from a colder body to a hotter one

# Second law: entropy

- Thermodynamics introduces a state function: **entropy**

$$\delta Q_{\text{rev}} = T\,dS$$

$$\oint \frac{dQ}{T} = \Delta S$$

$$\boxed{\begin{array}{cc} \Delta S \geq 0 & \text{for an isolated system} \\[2em] \Delta S = 0 \text{ (reversible)} & \Delta S > 0 \text{ (irreversible)} \end{array}}$$

# The laws of thermodynamics

**Minus first law** systems tend toward equilibrium

**Zeroth law** $A \sim B,\ B \sim C \Rightarrow A \sim C$

**First law** $dU = \delta Q + \delta W$

**Second law** $\Delta S \geq 0$

# Why is this philosophically puzzling?

- Why do isolated systems approach equilibrium?
- Why does entropy increase?
- Where does the asymmetry come from?
- How does thermodynamics relate to microphysics?

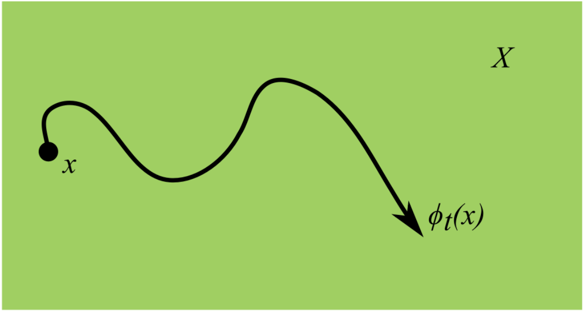
These are important questions: some argue *time asymmetry is grounded in thermodynamic asymmetry*

This is where statistical mechanics enters.

# Microscopic dynamics

- In classical mechanics, a system is represented by a state $x$ in a state space $X$
- Its evolution is given by a trajectory

$$x \mapsto \phi_t(x)$$



Thermodynamics must ultimately be grounded in this microscopic dynamics.

# The puzzle about the direction of time

- Thermodynamics is strongly **time-asymmetric**
- Isolated systems approach equilibrium
- Entropy increases:

$$\Delta S \geq 0$$

- But the underlying micro-laws are typically **time-reversal invariant**
- If a process is allowed, its time-reverse is allowed too

**No asymmetry in, no asymmetry out.**

Where does the thermodynamic arrow come from?

# Boltzmann’s first idea: the Boltzmann equation

- Focus on a **dilute gas**
- Describe it by a one-particle distribution $f_t(\mathbf{v})$
- $f_t(\mathbf{v})$: fraction of molecules with velocity **v** at time $t$

$$\frac{\partial f_t}{\partial t} = \text{collision term}$$

- Collisions gradually drive the gas toward the Maxwell–Boltzmann distribution

Idea: derive irreversible thermodynamic behaviour from molecular collisions.

# The $H$-theorem

- Boltzmann defines

$$H[f] = \int d^3v \, f(\mathbf{v}) \ln f(\mathbf{v})$$

- And proves, using the Boltzmann equation:

$$\frac{dH}{dt} \leq 0$$

- $H$ decreases until equilibrium
- Since $H$ behaves like **negative entropy**, this looks like:

$S$ increases until equilibrium

# Why this is problematic

- It works only under strong assumptions:
  - hard-sphere particles
  - dilute gas
  - only two-body collisions
  - **Stosszahlansatz / “collision number assumption"**
- More refined versions have been proposed...
- But molecular dynamics is time-reversal invariant:
  - this does **not** give genuine irreversibility
  - the theorem itself remains **time-reversal invariant**

So the *H*-theorem is not a straightforward derivation of the Second Law.

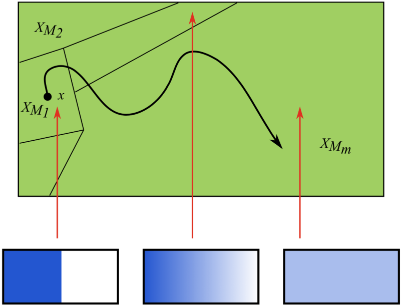
Boltzmann's first strategy is not enough.
This motivates a different framework:

**Boltzmannian statistical mechanics**

# Boltzmannian statistical mechanics: the basic idea

- A system has an exact **microstate** $x \in X$
- Thermodynamics describes it by a **macrostate** $M$
- Many different microstates can realise the same macrostate

$$X_M = \{x \in X : M(x) = M\}$$

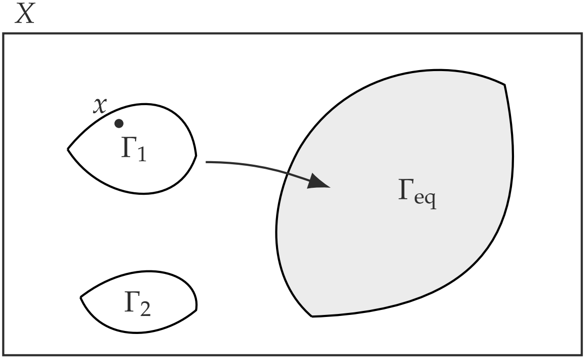


A macrostate corresponds to a **macro-region** of phase space.

## Boltzmann entropy

- Boltzmann associates entropy with the size of the macro-region

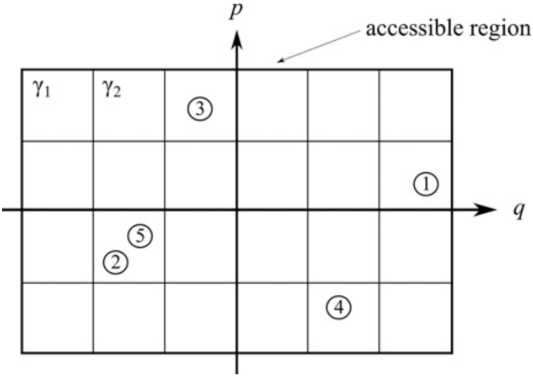
$$S_B(M) = k_B \log \mu(X_M)$$



Larger macro-region ⇒ larger entropy; equilibrium corresponds to the dominant region.

# Boltzmann’s combinatorial argument

- Start from the **one-particle space**
- Divide it into cells: **coarse-graining**
- A macrostate depends only on how many particles occupy each cell



The equilibrium distribution is the one compatible with the largest number of arrangements.

# Virtues and limits

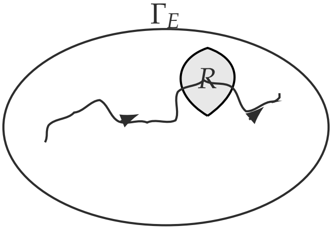
- It explains why higher entropy means: **more microscopic ways of being the same macroscopically**
- It connects equilibrium with overwhelmingly many compatible micro-arrangements
- But it has limits:
  - assumes non-interacting particles
  - depends on coarse-graining
  - does not yet explain *approach to* equilibrium

→ BSM needs an account of why systems move toward larger regions.

**Note**: *these approaches reject requirement of strict irreversibility!*

# First proposal: the ergodic explanation

- If the system is **ergodic** *(and many are not!)* its trajectory explores the whole accessible region
- In the long run: time spent in a region tracks that region's size

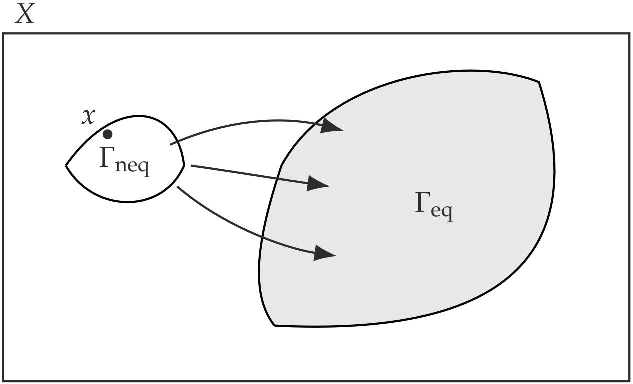


$$\text{fraction of time in } R \;\approx\; \frac{\mu(R)}{\mu(\Gamma_E)}$$

So if equilibrium is the biggest region, the system spends most of its time there.

# Second proposal: the typicality explanation

- Equilibrium occupies a **vast** region of phase space
- So equilibrium microstates are **typical**



Most microstates compatible with a non-equilibrium macrostate evolve into much larger macro-regions.

# Second proposal: the typicality explanation

> *"For a non-equilibrium phase point $x$ of energy $E$, the Hamiltonian dynamics governing the motion $x(t)$ would have to be ridiculously special to avoid reasonably quickly carrying $x(t)$ into [the equilibrium macro-region] and keeping it there for an extremely long time — unless, of course, $x$ itself were ridiculously special."*

Goldstein (2001, pp. 43–44)

- Objections: this is not enough. The dynamics could prevent the system from moving to the equilibrium macro-region.

# Third proposal: the Past Hypothesis

- Time-symmetric microphysics → if you explain why entropy is likely to increase toward the future from the system's present state, you get the mirror result to the past.
- *Idea*: Systems have higher chance of evolving to equilibrium and higher S because of (global) *initial conditions*:

"Given that the universe began in that special low-entropy state, what futures are typical from there?"

**Past Hypothesis (PH):**
the universe began in a very special
**low-entropy macrostate**

- Then, for **typical** microstates compatible with that beginning, entropy increases toward the future
- This is David Albert's **Mentaculus** picture.

# Past Hypothesis: philosophical questions

- What kind of posit is the Past Hypothesis?
  - a contingent initial condition?
  - or a **law-like** constraint on admissible histories?

- On a **Lewisian Best System** view, one natural question is:

  should PH belong in the best system of our world?
- Chen argues that PH is a **fundamental law**:
  - compatible with Humeanism
  - and also with a minimal non-Humean “governing” conception

- **Problems**:
  1. **Precision**: is PH well-defined?
  2. **Global vs. local**: how does a cosmic initial condition explain laboratory systems?
  3. **Dynamics**: why do typical states move toward higher S?
  4. **Quantum**: the behaviour is assumed classical/deterministic

# Where this leaves us

- **Ergodic strategy:** equilibrium is where the system spends most time
- **Typicality strategy:** equilibrium behaviour is overwhelmingly typical
- **Past Hypothesis:** the arrow of time comes from a special low-entropy beginning

None is entirely unproblematic.

Another framework is available:
**Gibbsian statistical mechanics**

# Gibbsian statistical mechanics

- Same phase space $X$ as in BSM, but a different idea:

  **not one microstate in one macro-region,**
  **but a probability distribution $\rho(x)$ on all of $X$**

- Macroscopic quantities are calculated as averages:

$$\langle f \rangle = \int_X f(x)\rho(x)\,dx$$

- And entropy is calculated from the distribution itself:

$$S_G = -k_B \int_X \rho(x) \ln \rho(x)\,dx$$

$S_G$ is high when $\rho$ is spread out fairly evenly over many microstates,
and low when it is sharply concentrated on a few.

# GSM: why useful? Why worry?

- **Why useful?**
  - it is the **main framework** for equilibrium calculations
  - it gives equations of state, averages, and fluctuations
- **Key difference from BSM**
  - **BSM**: one actual system has one actual microstate $x$
  - **GSM**: one system is described by a distribution $\rho$, often read via an **ensemble** of copies
- **Main worry**
  - how can an **ensemble average** describe one real system?
  - also: the Gibbs entropy $S_G$ is **constant** under the dynamics
  - ...so approach to equilibrium is harder to explain

So far, we mostly focused on one problem:

**Where does irreversibility of thermodynamics come from?**

→ We now turn to some other *major open issues* in the foundations of statistical mechanics.

# Open issues I: traditional topics

- **The BSM-GSM relationship**
  Incompatible/alternative? GSM more fundamental?
  (Wallace, 2018)

- **Probabilities in SM**
  What kind of probabilities are these?
  Not merely *subjective*: they are meant to track objective features of physical systems. Humean chances?
  (Loewer, 2001; Frigg and Hoefer, 2015)

- **Reduction of TD to SM**: **phase transitions**
  TD describes boiling or freezing as *sharp discontinuities*.
  SM recovers them exactly only with idealisations, notably the limit $N \to \infty$. Does this block reduction?
  (Palacios, 2019)

# Open issues II: hot topics

- **Relativity and thermodynamics**
  What becomes of temperature, pressure, and equilibrium in relativistic settings?
  (Chua, 2023; Wallace, 2025)
- **Quantum statistical mechanics**
  How should we understand equilibrium, entropy, and irreversibility when moving from CM to QM?
  (Shenker, 2022)
- **Black hole thermodynamics**
  Are black holes literally thermodynamic systems, with genuine temperature and entropy?
  (Robertson, 2025)

**Further references:**

- Frigg and Werndl (2024b) *Philosophy of Statistical Mechanics* (SEP)
- Frigg and Werndl (2024a) *Foundations of Statistical Mechanics*
- Uffink (2001) *Bluff Your Way in the Second Law of Thermodynamics*
- Wallace (2015) *The Quantitative Content of Statistical Mechanics*
- Myrvold (2021) *Thermodynamics: The science of heat and work*

**Textbooks:**

- Huang *Statistical Mechanics*
- Tong *Lectures on Statistical Physics*
- Mandl *Statistical Physics*

**Credits**: pictures 1, 2, 3 from Frigg and Werndl (2024b); picture 5 from Frigg and Werndl (2024a).

# References I

Chua, E. Y. S. (2023). T falls apart: On the status of classical temperature in relativity. *Philosophy of Science 90*(5), 1307–1319.

Frigg, R. and C. Hoefer (2015). The best humean system for statistical mechanics. *Erkenntnis 80*(3), 551–574.

Frigg, R. and C. Werndl (2024a). *Foundations of Statistical Mechanics*. Cambridge University Press.

Frigg, R. and C. Werndl (2024b). Philosophy of statistical mechanics. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2024 ed.). Metaphysics Research Lab, Stanford University.

Goldstein, S. (2001). Boltzmann's approach to statistical mechanics. In J. Bricmont and Others (Eds.), *Chance in Physics: Foundations and Perspectives*, pp. 39–54. Springer.

Loewer, B. (2001). Determinism and chance. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 32*(4), 609–620.

Myrvold, W. C. (2021). Thermodynamics: The science of heat and work. In W. C. Myrvold (Ed.), *Beyond Chance and Credence: A Theory of Hybrid Probabilities*, pp. 122–145. Oxford University Press.

# References II

Palacios, P. (2019). Phase transitions: A challenge for intertheoretic reduction? *Philosophy of Science 86*(4), 612–640.

Robertson, K. (2025). Black hole thermodynamics: A crisis of identity. *Australasian Journal of Philosophy*.

Shenker, O. (2022). Quantum foundations of statistical mechanics and thermodynamics. In E. Knox and A. Wilson (Eds.), *The Routledge Companion to Philosophy of Physics*, pp. Ch. 29. Routledge.

Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 32*(3), 305–394.

Wallace, D. (2015). The quantitative content of statistical mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 52*(Part B), 285–293.

Wallace, D. (2018, November). The necessity of gibbsian statistical mechanics.

Wallace, D. (2025). On relativistic thermodynamics. *Philosophy of Science 92*(3), 688–707.