Einstein's early work on Statistical Mechanics

A prelude to the *Marvelous Year*

Luca Peliti

M. A. and H. Chooljan Member
Simons Center for Systems Biology
Institute for Advanced Study
Princeton (USA)

November 6, 2015 / UFRGS Porto Alegre

Outline

Einstein before Einstein?

The 1902–04 Papers on the “Molecular Theory of Heat”

- Einstein’s approach to Statistical Mechanics is independent and bolder than Gibbs’
- Einstein focuses on fluctuations as a tool for discovery, rather than a nuisance
- The search for observable fluctuations leads him to focus on black-body radiation

The Papers

I. Kinetic theory of thermal equilibrium and of the second law of thermodynamics

6. *Kinetische Theorie des Wärmegleichgewichtes und des zweiten Hauptsatzes der Thermodynamik; von A. Einstein.*

II. A theory of the foundations of thermodynamics

9. *Eine Theorie der Grundlagen der Thermodynamik; von A. Einstein.*

III. On the general molecular theory of heat

6. *Zur allgemeinen molekularen Theorie der Wärme; von A. Einstein.*

Some biographical facts

- In 1902 Einstein had left the ETH having obtained a diploma in 1900, but not the doctorate
- In spring 1902 his application for Technical Assistant, 3rd Class, to the Federal Office for Intellectual Property in Bern was accepted, and he started working there in June
- He married Mileva Marić, whom she had met as a fellow student at ETH, in January 1903. Their first son was born in May 1904
- Before the three papers which interest us, he had published two papers in *Annalen der Physik*, which he much later judged “worthless”

Atomism in the XIX Century

- Chemists: Dalton, Avogadro, Cannizzaro
  The atomic idea becomes a *scientific tool*
- Early kinetic theory: Herapath, Waterston
  Forgotten for lack of observable consequences
- Kinetic theory: Clausius, Maxwell, Boltzmann, Loschmidt
  Maxwell: gas viscosity does not depend on density
  Connections with thermodynamics, the problem of entropy
- “Energetists” (e.g., Ostwald and Mach): Atoms are a concept and a calculating tool, not a reality
  (Cf. the position of the Church on Copernicanism)

The man who trusted atoms

- 1870: Ergodic hypothesis and physical interpretation of the temperature
- 1872: Boltzmann's equation and the *H*-theorem
- 1877: $S = k_{\mathrm{B}} \log W$ and the Boltzmann distribution for "complex molecules"
- 1884: Microcanonical and Canonical ensembles (respectively called *monode* and *holode*)

Einstein's motivations

- Einstein aims to "derive the postulates of thermal equilibrium and the second principle using exclusively the mechanical equations and the probability calculus"
- He provides "a generalization of the second principle, which is useful for the application of thermodynamics"
- He also gives the "mathematical expression of entropy from a mechanical point of view"
- The 1902–03 papers have similar structure: I'll deal with them in one go

Mechanical description

- General description of a mechanical system:

$$\frac{dp_i}{dt} = \phi_i(p_1, \ldots, p_n)$$

- Energy is the *unique* integral of motion:

$$E(p_1, \ldots, p_n) = \text{const.}$$

- (Liouville's theorem is only implicitly assumed):

$$\sum_i \frac{\partial \phi_i}{\partial p_i} = 0$$

Probabilistic description

- Observable quantities are given by temporal averages of functions of state variables:

$$\overline{A} = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt\, A(p_1(t), \ldots, p_n(t))$$

- For a given value of E, all observable quantities take on a constant value at equilibrium
- *Ergodic hypothesis*: for any region Γ of state space, let τ be the time spent in Γ during time $\mathcal{T}$. Then

$$\lim_{\mathcal{T}\to\infty} \frac{\tau}{\mathcal{T}} = \text{const.} = \int_\Gamma \epsilon(p_1, \ldots, p_n)\, dp_1 \cdots dp_n$$

Probabilistic description

- *Ensemble*: Given N systems of the same type, the number dN of systems in the small region g at any given time is

$$dN = \epsilon(p_1, \ldots, p_n) \int_g dp_1 \cdots dp_n$$

- From stationarity (and Liouville's theorem) one obtains

$$\epsilon(p_1, \ldots, p_n) = \text{const.}$$

- Einstein has thus derived the *microcanonical ensemble*

Canonical ensemble

- Consider a small system σ in contact with a much larger one Σ with total energy

$$E = \eta + H, \qquad E^* \leq E \leq E^* + \delta E^*$$

- Consider $g: \ \pi_i \leq \pi_i \leq \pi_i + \delta\pi_i$ $(= 1, \ldots, \ell)$ and $G: \ \Pi_i \leq \Pi_i \leq \Pi_i + \delta\Pi_i$ $(i = 1, \ldots \lambda)$:
- dN_1: number of systems that are found in $g \times G$:

$$\begin{aligned} dN_1 &= C \cdot d\pi_1 \cdots d\pi_\ell \, d\Pi_1 \cdots d\Pi_\lambda \\ &= \text{const.}\, e^{-2h(H+\eta)} d\pi_1 \cdots d\pi_\ell \, d\Pi_1 \cdots d\Pi_\lambda \end{aligned}$$

- Number of systems for which the state of σ lies in g:

$$dN \propto e^{-2h\eta} d\pi_1 \cdots d\pi_\ell \int_{E^*-\eta \leq H \leq E^*+\delta E^*-\eta} e^{-2hH} d\Pi_1 \cdots d\Pi_\lambda$$

Canonical ensemble

- $\chi(E) = \int_{E \leq H \leq E+\delta E} e^{-2hH} d\Pi_1 \cdots d\Pi_\lambda \simeq e^{-2hE}\,\omega(E)$
- Choosing h such that

  $$2h = \frac{\omega'(E)}{\omega(E)}$$

  χ is independent of the state of σ and we have

  $$dN = \text{const.}\ e^{-2h\eta}\, d\pi_1 \cdots d\pi_\ell$$

- The system σ acts like a thermometer, and if σ_1 and σ_2 are each in equilibrium with Σ, they are in equilibrium with each other ("0-th principle")
- Choosing σ as a single molecule, its average energy is equal to $3/4h$ and thus $2h = 1/k_{\mathrm{B}}T$ (in modern notation)

The Entropy

Infinitely slow processes

- Einstein considers two kinds of transformations:

  Adiabatic transformations: the evolution equations hold at every time, but the ϕ_i's can vary by external action via parameters λ

  "Isopycnic" (=equal-density) transformations: correspond to the thermal contact with a body at a different temperature: the evolution equations do not hold *during* the transformation, but before and after

- Any infinitely slow process can be approximated by a succession of adiabatic and isopycnic transformations

The Entropy

- During an infinitely slow process one has

$$dE = \sum \frac{\partial E}{\partial \lambda} d\lambda + \underbrace{\sum_{\nu} \frac{\partial E}{\partial p_\nu} dp_\nu}_{dQ}$$

- The canonical distribution holds before and after an infinitesimal transformation, thus from

$$dW = e^{c-2hE}\, dp_1 \cdots dp_n$$

one obtains from the normalization of W

$$\int e^{c+dc-2(h+dh)(E+\sum \partial_\lambda E\, d\lambda)}\, dp_1 \cdots dp_n = 0$$

(neglecting fluctuations in E) leading to

$$2h\, dQ = d\,(2hE - c)$$

The Entropy

Thus, since $1/4h = \kappa T$

$$dS = \frac{dQ}{T} = d\left(\frac{E}{T} - 2\kappa c\right)$$

leading to

$$\boxed{S = \frac{E}{T} + 2\kappa \log \int e^{-2hE}\, dp_1 \cdots dp_n}$$

But what about $\Delta S \geq 0$?

On the growth of entropy

Even Einstein has some difficulties with entropy growth...

- Consider an ensemble of N systems of energy between E and $E + \delta E$, and divide the available phase space into regions g_k of equal volume
- Define a "state" by assigning the number N_k of systems which lie in g_k
- Define the "probability" $\mathcal{W}$ of a state as the number of ways of distributing the systems compatible with the state. One has

$$\log \mathcal{W} = \log \frac{N!}{N_1! \cdots N_k! \cdots} \simeq \text{const.} - \int \rho_t \log \rho_t \, dp_1 \cdots dp_n$$

- Then $\mathcal{W}$ is maximal when the distribution is uniform

On the growth of entropy

- *"We have to assume"* that $\mathcal{W}$ never decreases: thus

$$-\int \rho_{t'} \log \rho_{t'} \, dp_1 \cdots dp_n \geq -\int \rho_t \log \rho_t \, dp_1 \cdots dp_n \quad \text{for } t' \geq t$$

- From this Einstein deduces (!) that $-\log \rho_{t'} \geq -\log \rho_t$ (again neglecting fluctuations...)
- Consider a collection of systems $\sigma_1, \sigma_2, \ldots$ initially isolated and let them exchange heat among themselves, then get isolated again and reach equilibrium
- The initial state $dw = dw_1 \cdot dw_2 \cdots = e^{\sum_\nu c_\nu - 2h_\nu E_\nu} \prod dp$ evolves into the final state $dw' = dw'_1 \cdot dw'_2 \cdots = e^{\sum_\nu c'_\nu - 2h'_\nu E'_\nu} \prod dp$
- Thus from $\rho_{t'} \leq \rho_t$ one obtains $\sum_\nu c'_\nu - 2h'_\nu E'_\nu \leq \sum_\nu c_\nu - 2h_\nu E_\nu$, i.e., $\sum_\nu S'_\nu \geq \sum_\nu S_\nu$

The 1904 Paper

On the general molecular theory of heat

- New expression from the entropy: given $\omega(E)\,\delta E = \int_{E<E(p)<E+\delta E} dp$ one has

$$S = \int \frac{dE}{T} = 2\kappa \int \frac{\omega'(E)}{\omega(E)} dE = 2\kappa \log[\omega(E)]$$

  $\omega(E)$ is a property of the *system*, not of the *environment*
- A new (more restricted) derivation of the second principle
- Interpretation of the constant κ: the average kinetic energy of a monoatomic gas at the temperature T is given by $3\kappa T$, yielding $\kappa = R/(2N_{\mathrm{A}}) = 6.5 \cdot 10^{-24}\,\mathrm{J/K}$ ($k_{\mathrm{B}} = 2\kappa \simeq 1.3 \cdot 10^{-23}\,\mathrm{J/K}$)

The 1904 Paper

On the general molecular theory of heat

Here the lion's paw starts to be felt...

- Einstein now considers fluctuations in E
- *The "general" meaning of* κ: from $\int dE\,(\overline{E} - E)e^{-E/2\kappa T}\omega(E) = 0$ one obtains

$$\overline{E^2} - \overline{E}^2 = \overline{\Delta E^2} = 2\kappa T^2 \frac{d\overline{E}}{dT}$$

- Application to radiation: where are the largest fluctuations expected? $d\overline{E}/dT$ is maximal when radiation intensity is maximal: $\overline{\Delta E^2} \simeq \overline{E}^2$
- But $\overline{E} = cvT^4$, then $\sqrt[3]{v} = 2\sqrt[3]{\kappa/c}/T \simeq 0.42/T\,\mathrm{cm}$ from Stefan-Boltzmann, while $\lambda_{\mathrm{max}} \simeq 0.293/T\,\mathrm{cm}$
- "This coincidence cannot be ascribed to chance, given the generality of our hypotheses"

Einstein vs. Gibbs

It is well known that while theory would assign to the [diatomic] gas six degrees of freedom per particle, in our experiments on specific heat we cannot account for more than five. Certainly, one is building on an insecure foundation, who rests his work on hypotheses concerning the constitution of matter.

Difficulties of this kind have deterred the author from attempting to explain the mysteries of nature, and have forced him to be contented with the more modest aim of deducing some of the more obvious propositions relating to the statistical branch of mechanics.

Gibbs, 1902, Preface

Einstein vs. Gibbs

Of special importance are the anomalies [fluctuations] of the energies, or their deviations from their average values. [. . .]

It follows that to human experience and observation [. . .], when the number of degrees of freedom is of such order of magnitude as the number of molecules in the bodies subject to our observation and experiment, $\varepsilon - \overline{\varepsilon}$, $\varepsilon_p - \overline{\varepsilon_p}$, $\varepsilon_q - \overline{\varepsilon_q}$, *would be in general vanishing quantities, since such experience would not be wide enough to embrace the more considerable divergencies from their mean values [. . .] In other words, such ensembles would appear to human observation as ensembles of uniform energy [. . .]*

Gibbs, 1902

Einstein vs. Gibbs

These relations [the "Jeans law"], found to be the conditions of dynamic equilibrium, not only fail to coincide with experiment, but also state that in our model there can be no talk of a definite energy distribution between ether and matter [. . .]

We therefore arrive at the conclusion: the greater the energy density and the wavelength of a radiation, the more useful do the theoretical principles we have employed turn out to be; for small wavelengths and small radiation intensities, however, these principles fail us completely.

In the following we shall consider the experimental facts concerning blackbody radiation without invoking a model for the emission and propagation of the radiation itself.

Einstein, 1905

Einstein vs. Gibbs

What brought Einstein to the blackbody problem in 1904 and to Planck in 1906 was the coherent development of a research program begun in 1902, a program so nearly independent of Planck's that it would almost certainly have led to the blackbody law even if Planck had never lived.

Kuhn, 1978

Summary

- Einstein considers thermodynamics as a perfect example of a “theory of principle”, starting from “empirically observed general properties of phenomena”
- He holds fast to the validity of the statistical principles, even in the presence of “insuperable difficulties”, but is ready to renounce aspects of Maxwell electrodynamics rather than the statistical principles (black-body radiation)
- He introduces (implicitly in 1904, explicitly in 1905) the “backward reading” of $S = k_{\mathrm{B}} \log W$
- In contrast with Gibbs, he welcomes fluctuations as a tool for investigating microscopic physics

Summary

- The first two papers of his “miraculous year” stem directly from his interest in fluctuations
- Looking for a “theory of principle”, analogous to thermodynamics, for electrodynamics led him to the Special Theory:

> *. . . we are by no means dealing with a ‘system’ here . . . but rather only with a principle which allows one to reduce certain laws to others, analogously to the second law of thermodynamics*

to Ehrenfest, 1909