Einstein and Boltzmann: Determinism and Probability

I. **History: from “Subtle is the Lord” by Abraham Pais**

In 1977, Boltzmann wrote a seminal paper:

“On the relation between the Second Law of Thermodynamics and Probability Theory with respect to the laws of thermal equilibrium.”

Three examples will be given of Einstein’s criticism of Boltzmann’s relation, which he called Boltzmann’s Principle:

\[ S = k \ln W. \]

Here \( W \) is the number of complexions, i.e. the number of microstates which correspond to a macrostate of the system.

Boltzmann used: \( N! / \prod_{j} n_j! \), where \( N \) is the number of particles in the system, \( n_j \) is the number of particles with a kinetic energy \( e_{\text{kin},j} = j \varepsilon \left( \varepsilon = \frac{1}{2} m v_j^2 \right) \).
1905  “The word probability is used in $S = k \ln W$ in a sense that does not conform to its definition as given in the theory of probability.”

1909  “[My] point of view is characterized by the fact that one should introduce the probability $[W]$ of a specific [macroscopic] state [of a system] in a *phenomenological* manner. In that way one has the advantage of not interposing any particular theory, for example, any Statistical Mechanics.”

This was done both by Einstein and Onsager, using the reverse expression: $W = e^{S/k}$, where $S$ is the entropy of the system in equilibrium or local equilibrium, respectively.

It was used for the theory of fluctuations of systems in equilibrium and near equilibrium (local equilibrium), respectively.
“Usually W is put equal to the number of complexions, “[However,] in order to calculate W, one needs a complete [deterministic] molecular-mechanical theory of the system under consideration. Therefore it is dubious whether the Boltzmann Principle has any meaning without a complete molecular-mechanical theory or some other theory which describes the elementary processes. S = k ln W seems [therefore] without content from a phenomenological point of view, without giving in addition such an ‘Elementary Theory.’”
I note that Einstein’s remark is applicable to all properties of a system in equilibrium and is an alternative to Gibbs’ entirely probabilistic approach to Statistical Mechanics (1902).

Therefore the question arises: Is a purely dynamical derivation of the equilibrium properties of a system possible?

I will discuss as an example the virial (density) expansions of the pair distribution function and the thermodynamic properties of a system in equilibrium.

I will do this by presenting first a formal virial expansion of the non-equilibrium pair distribution function and then reducing this to the virial expansion of the equilibrium pair distribution function.
The basic idea of virial expansions both in equilibrium and non-equilibrium

The virial (density) expansions reduce the [intractable] N-particle problem to a sum of [tractable] isolated groups of 2, 3, ..., s, ... particles, where each group is alone in the volume V of the system.

Density expansions will appear since the number of single particles, pairs of particles, ..., are proportional to n, n^2, ....

In equilibrium, for short-range particle interactions, this leads to virial (density) expansions of the thermodynamic properties (e.g. the pressure p) and the distribution functions, e.g. the pair distribution \( f_{2}^{e}(\tilde{q}_{1},\tilde{p}_{1},\tilde{q}_{2},\tilde{p}_{2} ; \beta) \), where \( \beta = \frac{1}{kT} \).

In non-equilibrium \( f_{2}^{ne}(\tilde{q}_{1},\tilde{p}_{1},\tilde{q}_{2},\tilde{p}_{2} ; t) \) leads to formal density expansions of the pair distribution function and of the dissipative transport coefficients, e.g. the shear-viscosity \( \eta \).
The virial expansion of the pair distribution function in equilibrium for short-range particle interactions

\[ f^e_2(x_1, x_2; \beta) = f^e_2(\tilde{q}_1, \tilde{p}_1, \tilde{q}_2, \tilde{p}_2; \beta) \]

\[ = n^e_2(\tilde{q}_1, \tilde{q}_2; \beta) \cdot f^e_1(p_1; \beta) f^e_1(p_2; \beta) \]

\[ n^e_2(\tilde{q}_1, \tilde{q}_2; \beta) = n^2 e^{-\beta \Phi_2(q_1, q_2)} + \]

\[ + n^3 \int d\tilde{q}_3 \left[ e^{-\beta \Phi_3(q_1, q_2, q_3)} - e^{-\beta \{\Phi_2(q_1, q_2) + \Phi_2(q_1, q_3)\}} - e^{-\beta \{\Phi_2(q_1, q_2) + \Phi_2(q_2, q_3)\}} + e^{-\beta \Phi_2(q_1, q_2)} \right] \]

\[ + O(n^4). \]

Here:

\[ \Phi_2(q_1, q_2) = \varphi(r_{12}) \]

\[ \Phi_3(q_1, q_2, q_3) = \varphi(r_{12}) + \varphi(r_{13}) + \varphi(r_{23}) \text{ (additivity)}; \ r_{ij} = |q_i - q_j| \]

\[ f^e_1(p; \beta) = ce^{-\beta p^2/2m}, \text{ where } c = (2\pi mkT)^{-3/2} \]
Ad the integrand of $n^3 \int d\vec{q}_3 \ldots$:

$$\left[ e^{-\beta \Phi_3(q_1,q_2,q_3)} - e^{-\beta \left( \Phi_2(q_1,q_2) + \Phi_2(q_1,q_3) \right)} - e^{-\beta \left( \Phi_2(q_1,q_2) + \Phi_2(q_2,q_3) \right)} + e^{-\beta \Phi_2(q_1,q_2)} \right]$$

This integrand has the *cluster property*, i.e., it vanishes for separated particle configurations, where any of the three particles does *not* interact with the other two.

Therefore, the integrand incorporates only *genuine* three-particle interactions or *overlaps*:

![Diagram of three particles](image_url)

This is necessary in order to avoid including again the contributions of (genuine) two-particle overlaps which have already been accounted for in the first two-particle term

$$e^{-\beta \Phi_2(q_1,q_2)}$$

This cluster property – to vanish for separated configurations – defines the nature of the clusters in this expansion. It was originally proposed by Ursell in 1927, who developed the general algorithm for this cluster expansion and distinguishes it from other cluster expansions.
The virial expansion of the pair distribution function in non-equilibrium

Non-equilibrium is fundamentally different from equilibrium in that dynamics has to be used. As a result, there is no separation of \( \bar{q} \) and \( \bar{p} \) as in equilibrium, since both are needed in dynamics.

In addition, the overlaps (“equilibrium collisions”) are replaced by genuine physical collisions between 2, 3, … isolated particles in the volume \( V \) of the system.
The density expansion of the *non-equilibrium* pair distribution function – i.e. the density to find two particles in the phases \(x_1 = \bar{q}_1, \bar{p}_1\) and \(x_2 = \bar{q}_2, \bar{p}_2\) – is for short-range forces:

\[
f_{2}^{\text{ne}}(x_1, x_2; t) = n^2 S_{-t}(x_1, x_2) f_{1}^{\text{ne}}(x_1; t) f_{1}^{\text{ne}}(x_2; t) + \]
\[+ n^3 \int dx_3 \left[ S_{-t}(x_1, x_2, x_3) - S_{-t}(x_1, x_2) S_{-t}(x_1, x_3) - S_{-t}(x_1, x_2) S_{-t}(x_2, x_3) + S_{-t}(x_1, x_2) \right].\]

and

\[
f_{1}^{\text{ne}}(x; t) f_{1}^{\text{ne}}(x; t) f_{1}^{\text{ne}}(x_3; t) + O(n^4).
\]

Here it has been assumed that \(f_{2}^{\text{ne}}(x_1, x_2; t)\) depends on the time \(t\) only via \(f_{1}^{\text{ne}}(x_1; t)\).
The streaming operator \( S_{-t}(x_1, ..., x_s) = e^{-t \mathcal{H}_s(x_1, ..., x_s)} \)
transforms the phases \( x_i \equiv (\bar{q}_i, \bar{p}_i) \) \( i = 1, ..., s \) of \( s \)-particles at
the initial time \( t = 0 \) into those at an earlier time \( -t \), using
Poisson’s form of Hamilton’s equations of motion, i.e., a
dynamical operator \( \mathcal{H}_s(x_1, ..., x_s) \), which solves the dynamical
\( s \)-particle problem.

Here \( \mathcal{H}_s(x_1, ..., x_s) = \sum_{i=1}^{s} \frac{\bar{p}_i}{m} \cdot \frac{\partial}{\partial \bar{q}_i} - \sum_{i < j}^{s} \theta_{ij} \) with
\[ \theta_{ij} = \frac{\partial \phi(r_{ij})}{\partial \bar{q}_i} \cdot \frac{\partial}{\partial \bar{p}_i} + \frac{\partial \phi(r_{ij})}{\partial \bar{q}_j} \cdot \frac{\partial}{\partial \bar{p}_j} \]
So that:
\[ \mathcal{H}_2(x_1, x_2) = \frac{\bar{p}_1}{m} \cdot \frac{\partial}{\partial \bar{q}_1} + \frac{\bar{p}_2}{m} \cdot \frac{\partial}{\partial \bar{q}_2} - \theta_{12} \]
\[ \mathcal{H}_3(x_1, x_2, x_3) = \frac{\bar{p}_1}{m} \cdot \frac{\partial}{\partial \bar{q}_1} + \frac{\bar{p}_2}{m} \cdot \frac{\partial}{\partial \bar{q}_2} + \frac{\bar{p}_3}{m} \cdot \frac{\partial}{\partial \bar{q}_3} - \left[ \theta_{12} + \theta_{13} + \theta_{23} \right] \]
The streaming operators in the integrand have the same
cluster property as in equilibrium, in that only genuine three-
particle collisions contribute to the integral.
Derivation of the equilibrium density expansion $f_2^e(x_1, x_2; \beta)$ from the nonequilibrium density expansion $f_2^{ne}(x_1, x_2; t)$

Nonequilibrium Virial Expansion

$$f_2^{ne}(x_1, x_2; t) = n^2 e^{-t \mathcal{H}_2(x_1, x_2)} f_1^{ne}(x_1; t) f_1^{ne}(x_2; t) + n^3 \int dx_3 \left[ e^{-t \mathcal{H}_1(x_1, x_2, x_3)} - e^{-t \mathcal{H}_2(x_1, x_2)} \cdot e^{-t \mathcal{H}_2(x_2, x_3)} \right] \cdot$$

$$\cdot f_1^{ne}(x_1; t) \cdot f_1^{ne}(x_2; t) \cdot f_1^{ne}(x_3; t) + \ldots$$

Equilibrium Virial Expansion

$$f_2^e(x_1, x_2; \beta) = n^2 e^{-\beta \Phi_2(\bar{q}_1, \bar{q}_2)} f_1^e(p_1; \beta) f_1^e(p_2; \beta) + n^3 \int dx_3 \left\{ e^{-\beta \Phi_3(\bar{q}_1, \bar{q}_2, \bar{q}_3)} + e^{-\beta \Phi_2(\bar{q}_1, q_2)} - \right.$$

$$\left. - e^{-\beta \Phi_2(q_1, q_2)} - e^{-\beta \Phi_2(q_1, q_3)} - e^{-\beta \Phi_2(q_2, q_3)} + e^{-\beta \Phi_2(q_1, q_2)} \right\} \cdot$$

$$\cdot f_1^e(p_1; \beta) f_1^e(p_2; \beta) f_1^e(p_3; \beta) + O(n^4)$$

Note that the first and second terms for genuine 2- and 3- particle overlaps or collisions can be written in Hamiltonian form:

$$H_2(x_1, x_2) = \Phi_2(q_1, q_2) + \frac{p_1^2}{2m} + \frac{p_2^2}{2m}; \quad H_3(x_1, x_2, x_3) = \Phi_3(q_1, q_2, q_3) + \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + \frac{p_3^2}{2m}$$
The equilibrium density expansion of $f_2^e(x_1, x_2; \beta)$ can be obtained from the nonequilibrium density expansion of $f_2^{ne}(x_1, x_2; t)$ by replacing $f_1^{ne}(\bar{q}_1, \bar{p}_1; t)$ by $f_1^e(p_1; \beta)$, the equilibrium Maxwell distribution of the particles’ velocity. One uses the following correspondence:

<table>
<thead>
<tr>
<th>Nonequilibrium</th>
<th>Equilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1^{ne}(\bar{q}_1, \bar{p}_1; t)$</td>
<td>$f_1^e(p_1; \beta)$</td>
</tr>
<tr>
<td>$e^{-t \mathcal{H}_s(x_1, \ldots, x_s)}$</td>
<td>$e^{-\beta H_s(x_1, \ldots, x_s)}$</td>
</tr>
<tr>
<td>$t$</td>
<td>$\beta = \frac{1}{kT}$</td>
</tr>
</tbody>
</table>
Term-by-term reduction of $f^{ne}_2(x_1, x_2; t) \rightarrow f^e_2(x_1, x_2; \beta)$

a. The crucial observation is that the Hamilton-Poisson operators $\mathcal{H}_s(x_1, \ldots, x_s)$ ($s = 2, 3, \ldots$) conserve the initial kinetic energy $\sum_{i=1}^{s} \frac{p_i^2}{2m}$ of the $s$ separated particles at the initial time, during the backward motion over time $-t$ for any time $t$.

b. Since only genuine three-particle collisions contribute to the three-particle integral, the initial kinetic energy transforms via $S_{-t}(x_1, x_2, x_3)$ into $H_3(x_1, x_2, x_3)$ at time $-t$.

c. One obtains then:

$$f^e_2(x_1, x_2; \beta) = n^2 e^{-\beta H_2(x_1, x_2)} + n^3 \int dx_3 e^{-\beta H_3(x_1, x_2, x_3)} + O(n^4),$$

where only genuine overlaps of the two or three particles are included.
For a thermodynamic property, like the pressure $p$, one obtains, using Clausius’ virial theorem:

$$
p = n k T \left[ 1 + \int d\bar{q}_2 n_2^c (\bar{q}_1, \bar{q}_2) \bar{q}_1 \cdot \frac{\partial \Phi_2 (\bar{q}_1, \bar{q}_2)}{\partial \bar{q}_1} \right],
$$

the virial expansion:

$$
p = n k T \left[ 1 + n \int d\bar{q}_2 \left\{ e^{-\beta \Phi_2 (\bar{q}_1, \bar{q}_2)} - 1 \right\} \right] + O(n^3).
$$

Here the cluster property follows for configurations, where $r_{12} = |\bar{q}_1 - \bar{q}_2| >$ range of the interparticle potential.
The divergences encountered in nonequilibrium are due to long collision sequences of binary collisions in time, and in equilibrium to long-range Coulomb forces between charged particles, as in an electron plasma or ionic solutions.

They are due to a basic physical flaw of the virial expansions in the presence of long-range memory effects in time or in space, due to the restriction to tractable small groups of particles in both cases and ignoring entirely collective, i.e. N-particle effects.

These divergences can be eliminated in nonequilibrium by considering tractable N-particle contributions by summing all most divergent collision sequences in each order of the density, leading to a collective mean-free time cutoff.
Similarly, in equilibrium, oppositely charged particles will arrange themselves by shielding each other, leading to an exponential decay $\sim e^{-r/\lambda_D}/r$, instead of $1/r$, of the Coulomb potential. Since both the mean-free time and the Debye shielding have density dependences, $\sim \frac{1}{n}$ and $\frac{1}{\sqrt{n}}$, respectively, the power series expansions in the density are replaced by non-analytic expressions in the density.
Conclusion

The dynamical nonequilibrium → static equilibrium derivation considered here for the virial expansions, accomplishes – for one non-trivial case – Einstein’s requirement of a dynamical, i.e. moving particles derivation, of equilibrium properties.
Or one could say, generalizing Galilei’s statement about the motion of the Earth, including the planets as well as the particles in a system in thermal equilibrium: “e pur si muovono”.