On the Mathematical Fractures at the Foundations of Statistical Mechanics

Dr. Roy Lisker
8Liberty Street#306
Middletown,CT 06457
rlisker@gmail.com
www.fermentmagazine.org

"Statistical Mechanics" is a reconstructive branch of physics that was created and developed by Ludwig Boltzmann (1844-1906) in a series of papers, articles and letters over the final third of the 19th century. It is unique in being a scientific theory designed to justify another scientific theory , a back-reconstruction of the kind of world that might exist at the microscopic level which would lead one to derive the laws of Thermodynamics from Hamiltonian Mechanics.

I*t is a brilliant achievement, even though all of it is based on sloppy, dubious or questionable mathematics !*

NOTE: This is meant to praise, not to condemn Ludwig:

In defense of Boltzmann's Questionable Mathematics

(1) Much of the mathematics Boltzmann needed was not discovered until much later: Probability, Topology, Set Theory, Measure Theory…

(2) Several fields in contemporary mathematics have been discovered through the "justifying constructions" of Statistical Mechanics: Ergodic Theory, Simplectic Geometry, Random Matrices, Markov Processes, etc.

(3) In many cases, Boltzmann's use of mathematics represents a kind of compromise between contradictory pictures of reality. This is fairly typical of physics in general.

(4) The inconsistent use of models for atoms, molecules, continua, collisions and so on, is not the fault of the mathematics, but of the pre-mathematical imagination of the physicist.

Boltzmann's ideas and innovations were expressed in many communications, and were summed up in the foundational papers of 1872, 1877 and 1887:

Foundational Papers of Statistical Mechanics

(1) 1872: *Further studies on the Thermal Equilibrium of Gas Molecules*

This paper introduces the famous, and still controversial, "H-Theorem", Boltzmann's attempt to prove the Second Law from elementary pictures of the sub-microscopic reality.

(2) 1877: In the communications of that year he defends the famous relationship between Entropy S, the Boltzmann constant k_B and phase-space "volume" W, $S=k_B\ln W$

(3) 1887: In response to the criticisms of Loschmidt , Zermelo and Poincaré concerning the irreversibility of the processes described by the H-Theorem, Boltzmann recast his 1872 paper into the language of probability. Thus "Statistical Mechanics" became truly statistical.

When I say that Boltzmann ought to be praised, not condemned for the expressionistic hand-waving that characterizes his mathematical demonstration, I have in mind a 6-phase model which I have developed for physical theories in general:

6 Phase Model for the interaction of Physics with Mathematics

Phase 1: **Careful observation of phenomena lead to the accumulation of "empirical data" for which there is no theory.**

Phase 2: **A model is proposed: thought experiments and mental pictures, concepts, hypotheses, laws.**

Phase 3: **Inconsistent mathematics (often incredibly so) is applied to the model to derive the data and formulae given in phase 1**

Phase 4: **Predictions are made from the mathematics; its imitations are exposed.**

Phase 5: **The mathematicians come in to fix the mathematical rigor. In the process they invent new mathematics**

Phase 6: **The process advances BOTH physics and mathematics.**

The fundamental paper of 1872:

This paper covers many topics. Of interest to us is a series of equations that go into the "H-Theorem", Boltzmann's attempt to derive the Clausius Inequality $\oint \dfrac{dQ}{T} \leq 0$ from elementary assumptions. As it turns out they are far from elementary. The expression under the integral sign is the negative of the Entropy. The closed circuit of the line integral refers to the *completion of a Carnot cycle*:

    a. Work is harnessed to produce heat

    b. Heat is transformed back into work

The Clausius Inequality expresses the monotonic increase of Entropy, and is equivalent to the Second Law: In any such cycle the "work" (Free available energy) returned is less than the amount invested. "Equality" is only a limit condition.

The interpretation of the formula for H is interesting. This is not the negative of the entropy, *but a quantity which evolves to the entropy as the system evolves over time to a stable equilibrium temperature.* Thus, the H-theorem could, in theory, be used to characterize all states, gaseous, liquid and sold.

However, the equations for the H-Theorem have only been solved for special cases, and are only useful for dilute gases in equilibrium, as one finds in aeronautics. Solving the H-Theorem under particular assumptions is a thriving branch of modern mathematics.

Boltzmann's own demonstration of the H-Theorem and its connection to the Second Law, is based on 3 assumptions, all of them questionable when not absurd:

(1) *The Equipartition Theorem:* **For a dilute gas in a compact region in equilibrium, the locations of its molecules (atoms, electrons etc.) on any constant energy surface in phase-space are uniformly distributed.**

(2) The *Stoss-Zahl- Ansatz* **(Collision Number Hypothesis, or Hypothesis of Molecular Chaos). This has two parts:**

(a) The velocities of colliding molecules are uncorrelated before collision, but correlated after collision. **Boltzmann needs this to derive Irreversibility . Unfortunately, as was pointed out by Johan Loschmidt, it already contains irreversibility in its definition.**

(b) The collision locations are uniformly distributed both before and after collisions. This is not explicitly stated but is evident in the way the equations are interpreted.

(3) *The Ergodic Hypothesis.* **Apart from a statistically insignificant percentage the states (x, v) of each particle permute through every cell of a discretized phase space volume occupied by the system. Essentially, the trajectory will be dense in the compact phase space volume. The combination of the Stoss-Zahl-Ansatz and Ergodic Hypothesis**

Elliot Lieb (Paraphrase)

Statistical Mechanics is based on 3 absurd notions:

(1)*The Ergodic Hypothesis is ridiculous*

(2)*The Equipartition assumption is ad hoc*

(3)*The Stoss-Zahl-Ansatz is self-contradictory*

Why, then, does Statistical Mechanics work?

(1) To justify the 3 assumptions , Boltzmann assembled a series of

pictures of the entities in the sub-microscopic reality:

   a. Molecules are hard, perfectly elastic spheres.

   b. Were the "state, $(x, dx/dt, t)$" knowable with perfect

accuracy,these spheres could be could be reduced to points. Then one

could apply ordinary Hamiltonian mechanics

   c. Since this is not possible, one has to invoke a local

probability distribution function $f = f(x,v,t)$, and a global probability ρ

gotten from integrating this over phase space.

Over the years, Boltzmann flounders (*Republican might say "flip-flops"*) between these and other images, all designed to reify his intuition of an invisible atomic and molecular world. The Atomic Hypothesis was not generally accepted in his day, and condemned by the influential authority of Wilhelm Ostwald, Ernest Mach and others adhering to the Positivist approach to science. In fact, it was *through* Statistical Mechanics that Einstein explained and quantified Brownian Motion. Building on Einstein's work, atoms and molecules were detected by Jacques Perrin in 1912.

Summarizing: To justify his probability distribution densities for discrete particles, Boltzmann reduces the hard spheres to points. He is then obliged to chop up, or discretize , the phase space into tiny boxes, which in theory can be counted. ("coarse-graining"). They have a limiting size that was later determined from Planck's constant, h.

Using this procedure he defines a "fictive density" of an effectively *massless* fluid. More precisely, all the molecules are assume to have *the same mass*, which cannot be compressed into larger and smaller densities. However one can "count" the number of boxes through which an individual molecule or a system happens to pass. Eventually, Kolmogorov will replace "counting" with "Lebesgue measure".

Thus, by invoking probabilities, (and Boltzmann will eventually uses4 different notions of probability) , he si able to distinguish

between "typical" and "rare" configurations of a confined gas. This involves further assumptions of the passage from microstates to macrostates which we won't go into.

(2) *The hypothesis of molecular chaos:"Stoss-Zahl-Ansatz")(SZA) .*

This has been criticized more extensively than any other assumption underlying Statistical Mechanics. No one is happy with the Stoss-Zahl-Ansatz. In 1912, Paul and Tatania Ehrenfest developed "toy models" ,dubbed the "Wind-Tree Model" and the "Dog-Flea Model" that attempts to reproduce the consequences of the assumptions of molecular chaos. There have been several other such models , including the "Kac-ring model". These games are too unrealistic to be applicable, although they do indicate tendencies confirming molecular chaos.

Boltzmann invokes the SZA when he needs to justify the irreversibility of the 2nd Law, then discards it when it proves inconvenient to the probabilistic approach. Simply stated, he assumes that the probabilities of joint distributions are equal to the simple products of the individual distributions, whether or not there is independence or correlation.

(2) The Ergodic Hypothesis. This has several variants. In its strongest form it appears in the paper of 1872 to justify the Equipartition Theorem: As stated by Boltzmann, it means that almost all particles, (molecules, points, probability densities) of a thermodynamic system in

equilibrium, will past through every discrete cell in the phase space of the system before returning to its original location.

As employed today, the Ergodic Hypothesis states that, in a system of microstates, the amount of time that the sub-systems of particles of equal energy, will remains in a given region for a length of time proportional to its energy. Boltzmann periodically(!) uses and abandons the Ergodic Hypothesis. Like String Theory it has generated (as has most of Statistical Mechanics) a rich branch of mathematics.

Summarizing the dubious mathematics:

(1) The Ergodic Hypothesis

(2) The Equipartition "Theorem"

(3) Stoss-Zahl-Ansatz (molecular chaos)

(4) Treating the micro-canonical ensemble as a continuum

(5) Discretization of phase space

(6) Reversion to a continuum model for the phase space, to get the Maxwell-Boltzmann distribution

(7) The fictive density of states

(8) Criterion for "equivalence" of microstates in the macrostate

(9) Loschmidt , Poincare and Zermelo objections.

Criticisms of Loschmidt, Zermelo and Poincare

These are merely the most famous names in a controversy that raged in the scientific journals of Germany, Austria, France and England at the end of the 19th century. In fact, all of them expressed numerous criticisms, of which these are a selection:

Johan Joseph Loschmidt's Objection: A Hamiltonian Dynamics based entirely on reversible processes cannot serve as the basis for an irreversible law

Ernst Zermelo's Objection: There is no evidence that the critical point for integral expression for the entropy in the H-Theorem is stable, i.e., that one can even talk about an equilibrium state that is equated with the temperature

Henri Poincaré's Objection: The Poincaré Recurrence Theorem , the basis of Mathematical Ergodic Theory: any measure preserving function on a compact manifold will come arbitrarily close to its initial state infinitely many times.

Deriving the Thermodynamic Equations from Combinatorics.

All Thermodynamic quantities *for a system in equilibrium* can be derived from simple combinatorics and lots of hand-waving. It does not supply a collision mechanism to explain Entropy and the Second Law, but it does lead to the same result, that dH/dt $\leq$ 0. One indeed wonders if the H-Theorem is needed at all!

The contribution of Boltzmann was:

(1) An independent derivation of the Maxwell's Gaussian distribution of energies for a gas in equilibrium

(2) The Clausius Inequality and Second Law, that is to say, the phenomenon of "dissipation"

(3) Support for the atomic hypothesis, leading eventually to the detection of atoms in 1912

(4) The so-called "Boltzmann equation" which equates Entropy with the logarithm of phase space volume.

(5) H is defined for both equilibrium and non-equilibrium states.

An excellent presentation of this hand-waving appears in Erwin Schrodinger's "Statistical Thermodynamics" (Dover). I want to briefly go over this with you to show

(1) How easily this is done

(2) How cavalier the attitudes of even the best theoretical physicists are towards mathematical rigor

(3) To lay the basis for a presentation of Boltzmann's H-Theorem.

Consider N identical systems, let's say electrons. The list of possible energy eigenvalues is

$$\boxed{\begin{array}{l} \varepsilon_1, \varepsilon_2, ..., \varepsilon_l, ... \\ \varepsilon_{l+1} \geq \varepsilon_l \end{array}}$$

If the system is classical, it is completely determined once one knows that S_1 is in state l_1 , S_2 in state l_2, etc. Each state has an "occupation number", a_j , which gives the number of electrons in a given state.

$$\boxed{\begin{array}{l} P = \dfrac{N!}{a_1! a_2! ... a_l! ...} \\[2mm] \sum a_l = N; \ \sum a_i \varepsilon_i = E \\[2mm] \dfrac{E}{N} = \dfrac{\sum a_i \varepsilon_i}{\sum a_l} \end{array}}$$

Boltzmann observed that the maximum for P is astronomically much larger than any lesser value. This can be shown rigorously for $N \to \infty$. However, when N becomes small one must pay

attention to the fluctuations of Brownian motion. The standard treatment
now follows. To maximize *ln* P, we use the technique of Lagrange
multipliers to the expression :

$$A = \ln P - \lambda N - \mu E = \ln N! - (\sum \ln a_i!) - \lambda(\sum a_i) - \mu(\sum \varepsilon_i a_i)$$

All of the a_i's are treated as if they were continuous,
independent variables, although in fact they are integers. A better
approach therefore would be to use the ratios of the a_i's to N, which, in
the limit can be approximated as continuous variables. In any case, one
invokes a crude approximation of Stirling's Formula $\boxed{\ln k! \approx k(\ln k - 1)}$

Taking the derivative:

$$A \approx \ln N! - \sum a_l(\ln a_l - 1) - \lambda(\sum a_i) - \mu(\sum \varepsilon_i a_i)$$
$$\frac{\partial A}{\partial a_l} \approx -\ln a_l + 1 - 1 - \lambda - \mu \varepsilon_i$$

Hence

$$\boxed{\ln a_l + \lambda + \mu \varepsilon_l = 0}$$

Solving for each a_i gives :

$$a_l = e^{-\lambda - \mu \varepsilon_l}$$
$$N = \sum e^{-\lambda - \mu e_l}$$
$$E = \sum \varepsilon_l e^{-\lambda - \mu e_l}$$

These are the basic formulae of of Statistical Mechanics,

from which the quantities of Thermodynamics, and the partition function, are derived. The partition function is the expression for E/N.

Commentary

Mathematically this procedure is outrageous!! The a_j's are huge integers, one can hardly speak of "differentiating". In the Stirling formula, these extremely discrete functions making huge leaps are replaced by an "asymptotically continuous" function, with meaningless infinitesimal increments producing controllable increases on the range.

Worse still, one applies Langrangian Multipliers! Obviously there are other means to the same results, but the procedure is absurd!

Continuing:

$$U = \frac{E}{N} = \frac{\sum \varepsilon_l e^{-\mu e_l}}{\sum e^{-\mu e_l}}$$

$$= - \frac{\partial(\ln(\sum e^{-\mu e_l}))}{\partial \mu}$$

Eliminating λ :

$$\lambda = -\ln a_l - \mu\varepsilon_l$$

The multiplier μ turns out to be the inverse of the Boltzmann constant times the temperature. This is not surprising as this inverse is the integrating factor for dQ.

The basic equation of thermodynamics is dS = 1/T(dU +pdV). The equivalent equation in Schrodinger's derivation is:

$$\mu = F(T) = \frac{1}{kT}$$

$$d(F + U\mu) = \mu(dU - \frac{1}{N}\sum q_l d\varepsilon_l)$$

Where F is the "free energy". Substituting

in previous formulae gives the classical form of the Partition function:

$$Z = \sum e^{-\frac{e_l}{kT}}$$

The free energy is given by

$$F(T) = -k \ln Z = \frac{U - ST}{T}$$

****

An Outline of the H-Theorem.

This is a simplified and modernized treatment combining the demonstrations of Cedric Villani and Harvey Brown in the Bibliography.

H is a quantity that behaves *like* the Entropy, for which reason Boltzmann assumes that it *is* , in essence, the expression for the Entropy. This is very much questioned today. Boltzmann's arguments rather show that the convergence of the Entropy to an equilibrium temperature under the effect of dissipation is stable.

Properties of the "H" function

(1) H is a function of the state variables: x, t, v , T and probability density ρ. (x and v are of course abbreviations for 3-tuples (x,yz) and (v_x, v_y, v_z))

(2) The definition of H depends on the Transport Operator D (explained below). This is where all the trouble is, particularly with respect to the Stoss-Zahl-Ansatz

(3) Its' domain is a compact region of phase space, over an indefinitely long period of time

(4) Given the 3 assumptions cited by Elliot Lieb, combined with the Liouville Theorem that asserts the invariance of the volume of phase space covered by an ensemble of systems as it moves through time under a Hamiltonian flow, Boltzmann derives the result dH/dt $\leq$ 0 for all time.

Guiding the methodology is the dogma that all measurable macroscopic quantities can be derived from microscopic averages. Let **f** be the probability distribution of a dilute gas confined to a compact region. Then the probability of the system being in a given state is

$$\rho = \int_{R^3} f(t,x,v)\,dxdtdv$$

If x and t are fixed, this becomes an integral of the single variable ,v.

From Liouville's Theorem we know that this "density" propagates without compression, expansion or change. (The total time derivative of f can be broken into a temporal part and a gradient:

$$D_t f = \frac{\partial f}{\partial t} + v \bullet \nabla f = 0$$

The gradient part is known as the "transport operator". If there is a macroscopic force, **F**, then this equation is given a Newtonian modification as:

$$D_t f = \frac{\partial f}{\partial t} + v \bullet \nabla_x f + F \bullet \nabla_v f = 0$$

The 6 assumptions:

(1) The gas is dilute

(2) Collisions are very brief events at very precise locations ,*x* .

(3)   Collisions are assumed perfectly elastic. Therefore, given:

Velocities before collision $v^{'}, v_{*}^{'}$

Velocities after collision v, v_{*}

$$v^2 + v_{*}^{2} = v'^{2} + v_{*}'^{2}$$

$$v + v_{*} = v' + v_{*}'$$

Then $v^{'} = \dfrac{v + v_{*}}{2} + \dfrac{|v - v_{*}|}{2}\sigma$

$$v_{*}' = \dfrac{v + v_{*}}{2} - \dfrac{|v - v_{*}|}{2}\sigma$$

σ is the "deflection angle", σ = sinθ

(4) Everything is microreversible

(5) Molecular chaos (Stoss-Zahl-Ansatz!) "The velocities of particles before collision are uncorrelated"

(6.) Only binary collisions are considered. The percentage of multiple collisions is assumed to be negligible.

Using these 6 assumptions, Boltzmann derived the Quadratic Collision Operator, which we write out in full:

$$v^2 \frac{\partial f}{\partial t} = D_t f - v \bullet \nabla_x f = Q(f,f)$$

$$= \int_{R^3} dv_* \int_{S^2} B(v - v_*, \sigma)(f' f'_* - f f_*) d\sigma$$

The "flux expression" $J = f' f'_* - f f_*$ represents the *joint probability distribution of velocities before collision minus the joint probability distribution after collision*. That these are simple products is a consequence of the hypothesis of molecular chaos.

The form of the kernel B depends on the model employed for the physical image of the particle: hard sphere, point particle, field, linearized field, etc. In general the kernel, is not integrable.

The flux expression J in f is a tensor product in probabilities, which is allowed because the particles are uncorrelated before collision. Even though they are no longer uncorrelated after collision, Boltzmann continues to use the same expression. This in essence is the Loschmidt objection.

Here is how the 6 assumptions go into the integral and the theorem:

(1) t and x are treated as parameters, that is to say, the collisions are localized in time and space

(2) Collisions are perfectly elastic, as required for the tensor product

(3) The microreversibility is built into the structure of the kernel B

(4) Stoss-Zahl-Ansatz

Note that Df is linear, while Q(f,f) is non-linear. Boltzmann then constructs the following integrals

$$Q(f,f) = \int_{R^3} dv_* \int_{S^2} B(v - v_*, \sigma)(f'f_*' - ff_*)d\sigma$$

$$\int Q(f,f)\psi dv = \int dv \int_{R^3} dv_* \int_{S^2} B(v - v_*, \sigma)(f'f_*' - ff_*)\psi d\sigma$$

$$\psi = \ln f$$

$$\int Q(f,f)\ln f dv = -Df$$

The choice of lnf for ψ comes from the combinatorial arguments outlined above. Following an (excessive!) series of manipulations involving changes of variables and integration by parts, which takes up many pages and could certainly have been simplified, Boltzmann arrives at:

$$\int D(f)dt =$$

$$\tfrac{1}{4}\int\int\int dv dv_* d\sigma B(v - v_*)(f'f_*' - ff_*)\ln(\frac{f'f_*'}{ff^*})$$

Note that the integrand is of the form (X-Y)(lnX-lnY). Assuming that the kernel is positive, this means that the integral will always be > 0. Therefore the derivative will be positive, and the quantity D will always be increasing.

A Short List of Criticisms to the H-Theorem

(1) Boltzmann Identifies D with the entropy;

(2) Assumes that it rises to a maximum;

(3) Assumes that, once this maximum is reached, it will remain constant for all time.

(4) That the progress to the critical point will happen in finite time;

(5) Assumes that the integrand is continuous.

(6) Assumes that this maximum is stable, that is, there will not be jumps along the way;

(7) Assumes that the maximum is finite;

(8) Argues that the distribution at the unique critical point will be the Maxwell-Boltzmann distribution derived from the equipartition of energies.

Bibliography

(1)   Harvey R. Brown, Wayne Myrvold: *Boltzmann's H-Theorem, its limitations and the birth of (fully) statistic mechanics.*
www.philsci-archive.pitt.edu/4187/

(2)   Jos Uffink: *The Boltzmann Equation and H-Theorem*
www.pitp.phas.ubc.ca/confs/7pines2009/readings/Uffink.pdf

(3) Sergio B. Volchan : *Probability as typicality*
http://arxiv.org/abs/physics/0611172

(4) Cedric Villani: *Mathematical topics in collisional kinetic theory* Handbook of Mathematical Fluid Dynamics (Vol. 1), edited by S. Friedlander and D. Serre, published by Elsevier Science (2002)

(5) Vincent S. Steckline: *Zermelo, Boltzmann and the recurrence paradox* Am. J. Physics 51 (10) October 1983

(6) Ilya Progogine: *From Being to Becoming* WH Freeman 1981

(7) Stephen Brush: *The Kinetic Theory of Gases: An Anthology of Classic Physics* Imperial College Press 2003

(8)

(9) Enrico Fermi: *Thermodynamics* Dover, 1936

(10) Gallavotti, Reiter, and Yngvason editors: *Boltzmann's Legacy* European Mathematical Society, c2008.

(11) A d'Abro: "The Rise of the New Physics" Vol 1, Thermodynamics; Kinetic Theory Dover 1939

(12) Ehrenfest, Paul and Tatiana: Conceptual Foundations of the Statistical Approach in Mechanics Cornell University Press 1959

(13) E Schrödinger: "Statistical Thermodynamics" Cambridge1964

(14) Bergmann, PG "Basic Theories of Physics: Heat and Quanta" Dover 1951

(15) Georg Joos: "Theoretical Physics", translated by Ira M. Freeman Hafner Publishing 1934 Part V, Theory of Heat ; Part VI Theory of Heat, Statistical Part