The paradox of irreversibility

Oliver Penrose

Department of Mathematics and
the Maxwell Institute for Mathematical Sciences,
Heriot-Watt University, Edinburgh

Talk to Mathematics Department
University of Bath, Friday 27 Jan 2017

Outline

- **The paradox of irreversibility**
- A. Stating the paradox
- B. Boltzmann's work and his critics
- C. Some quotations
- D. Explanation versus description
- E. Identifying "good" motions

1. The paradox of irreversibility

- Matter is made out of particles (atoms)
- A primary aim of statistical mechanics is to derive the behaviour of matter from the laws desribing the motion of these particles
- BUT
- The differential equations (‘laws of microscopic mechanics’) describing the motion of these particles are symmetrical under time reversal
- The actual behaviour of matter is not symmetrical under time reversal
- AND YET
- The actual behaviour of matter is supposed to be derivable from these laws

2. Not all mechanical laws of motion conserve energy

- Example of a (1-D) mechanical law with no friction

$$m\,d^2x/dt^2 = -kx \quad (\text{with } k > 0)$$

where $x(t)$ denotes position of particle at time t.

- This law conserves energy because

$$\frac{d}{dt}\left(\tfrac{1}{2}m\left(\frac{dx}{dt}\right)^2 + \tfrac{1}{2}kx^2\right) = \left(m\frac{d^2x}{dt^2} + kx\right)\frac{dx}{dt} = 0$$

- Example of a mechanical law with friction

$$md^2x/dt^2 = -kx - \alpha dx/dt \quad (\text{with } k, \alpha > 0)$$

- This law doesn't conserve energy :

$$\frac{d}{dt}\left(\tfrac{1}{2}m\left(\frac{dx}{dt}\right)^2 + \tfrac{1}{2}kx^2\right) = \left(m\frac{d^2x}{dt^2} + kx\right)\frac{dx}{dt} = -\alpha\left(\frac{dx}{dt}\right)^2 \le 0$$

3. Equations of motion for a system of N particles

- In statistical mechanics we can represent a material system microsopically as a system particles with eqns of motion of the form

  $$m\frac{d^2x_i}{dt^2} = -\frac{\partial V(x_1,\ldots,x_n)}{\partial x_i} \quad (i = 1,\ldots,n)$$

  where $x_1 \ldots x_n$ are the position coordinates, $n(= 3N)$ is the (very large) number of degrees of freedom, and $V(x_1,\ldots,x_n)$ is a given function, the potential energy.
- This eqn of motion does conserve energy:

  $$\frac{d}{dt}\left(\frac{1}{2}m\sum_{i=1}^{n}\left(\frac{dx_i}{dt}\right)^2 + V\right) = \sum_{i=1}^{n}\left(m\frac{d^2x}{dt^2} + \frac{\partial V}{\partial x_i}\right)\frac{dx_i}{dt} = 0$$

- This eqn of motion contains no frictional terms proportional to velocities such as dx_i/dt. If friction is present, it does not alter the total energy; rather it converts useful mechanical energy into heat energy but the total energy is conserved.

4. An example of time-reversal symmetry

- for the simple harmonic oscillator with eqn of motion

  $$d^2x/dt^2 = -x$$

  $x(t) = \sin(t)$ is a solution of the eqn of motion;
  $x(t) = \sin(-t)$ is another.
- A more general oscillator, whose eqn of motion is

  $$m\frac{d^2x}{dt^2} = -\frac{dV(x)}{dx}$$

  If $x = \xi(t)$ is a solution, i.e. if $m\xi''(t) = -V'(\xi(t))$ for all t, then $x = \xi(-t)$ is also a solution.
- The solutions have this *time-reversal symmetry* because the DE is invarant under the transformation $t \to -t$
- An eqn of motion with frictional terms, e.g.

  $$\frac{d^2x}{dt^2} = -\alpha\frac{dx}{dt} - kx \quad (\alpha > 0)$$

  does *not* have time-reversal symmetry

Most macroscopic PDE's are not time-symmetric

example: the PDE for the temperature $\theta(\mathbf{x}, t)$ at position $\mathbf{x}$, time t in a uniform isotropic solid is

$$\frac{\partial\theta}{\partial t} = D\nabla^2\theta$$

This is not time-reversal symmetric, nor are its solutions

Kelvin's description of a (notional) time-reversed motion

"If . . . the motion of every particle of matter in the universe were precisely reversed at any instant the course of nature would be simply reversed for ever after. The bursting bubble of foam at the foot of a waterfall would reunite and descend into the water: the thermal motions would reconcentrate their energy and throw the mass up the fall in drops re-forming into a close column of ascending water. Heat which had been generated by the friction of solids and dissipated by conduction, and radiation with absorption, would come again to the place of contact, and throw the moving body back against the force to which it had previously yielded. Boulders would recover from the mud the materials required to rebuild them into their previous jagged forms, and would become reunited to the mountain peak from which they had formerly broken away." (W. Thomson, *The kinetic theory of the dissipation of energy,* Proc. Roy. Soc. Edinb., 1874)

B. Boltzmann's integro-differential equation

- For a low-density gas of N hard spheres with radius a confined to a region $\Lambda \in \mathcal{R}^3$ define the one-particle distribution function $f(\mathbf{x}, \mathbf{v}, t)$ so that

  $$\text{Number of particles in } \Delta \text{ at time } t \approx \int_{\Delta} d^3\mathbf{x} \int_{\mathcal{R}^3} d^3\mathbf{v} f(\mathbf{x}, \mathbf{v}, t)$$

  where Δ is any "box" in the phase space $\Lambda^N \mathcal{R}^{3N}$, large enough to contain a huge number of particles.
- Boltzmann's eqn is

  $$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = \left(\frac{\partial f}{\partial t} \right)_{coll'n}$$

  where the last term represents the rate of change of f due to collisions (see next frame).

Boltzmann's collision formula for a gas of hard spheres

- $$\left(\frac{\partial f}{\partial t}\right)_{coll'n} = N \int_{\mathcal{R}^3} d^3\mathbf{v}_1 \int_{\mathbf{n}\ :\ \mathbf{n}\cdot(\mathbf{v}-\mathbf{v}_1)>0} d^2\mathbf{n}\ \times$$

  $$\times a^2 \mathbf{n}\cdot(\mathbf{v}-\mathbf{v}_1)[f(\mathbf{x},\mathbf{v}_1')f(\mathbf{x},\mathbf{v}') - f(\mathbf{x},\mathbf{v})f(\mathbf{x},\mathbf{v}')]$$

  where $\mathbf{v}', \mathbf{v}_1'$ are the post-collision velocities of two particles whose pre-collision velocities are $\mathbf{v}, \mathbf{v}_1$, the unit vector along their line of centres at the moment of collision being $\mathbf{n}$.
- It was proved by Lanford (1975), subject to certain plausible assumptions about the initial conditions, that Boltzmann's integro-differential equation holds exactly in the limit where $N \to \infty, Na^2 \to \mathrm{const.}$, for a limited time after the initial time (unfortunately, only about $1/5$ of the mean time between collisions)

Boltzmann's "H" theorem

- For a gas in a container Λ obeying Boltzmann's integro-differential equation the quantity

  $$H(t) := \int_\Lambda d^3\mathbf{x} \int d^3\mathbf{v} f(\mathbf{x},\mathbf{v},t) \log f(\mathbf{x},\mathbf{v},t)$$

  satisfies

  $$\frac{dH(t)}{dt} \leq 0$$

- Since H is bounded below, it approaches a limit. The limiting (equilibrium) $f(\mathbf{x},\mathbf{v},t)$ is the t-independent Maxwellian distribution

  $$f(\mathbf{x},\mathbf{v},t) \propto \exp\left(-\beta \tfrac{1}{2} m |\mathbf{v}|^2\right)$$

  where m is the mass of each particle and β is a constant related to the temperature.

- Physical interpretation: the (non-equilibrium) entropy S is related to H by $S = -kH$. The constant k is called Boltzmann's constant)

Loschmidt's 'reversibility' objection to Boltzmann's theory

- The H theorem, $dH(t)/dt \leq 0$ is not reversible. How can it follow from the reversible equations of Newtonian dynamics?
- Boltzmann: "My minimum theorem as well as the so-called Second Law of Thermodynamics are only theorems of probability.It can never be proved from the equations of motion alone that H must always decrease. It can only be deduced from the laws of probability, that if the initial state is not specially arranged for a certain purpose, but [chance] governs freely, the probability that H decreases is always greater than that it increases." (*Nature* **51** 413-5 (1895))
- Boltzmann again: In general when a gas . . . has initially some improbable state, then it passes to the most probable state . . . and remains there . . . The one-sidedness of this process is clearly not based on the equations of motion of the molecules. For these do not change when the time changes its sign. This one-sidedness rather lies uniquely and solely in the initial conditions." (pp 441-442 of *"Lectures on Gas Theory" (1896)* translated by S G Brush (Univ. of California Press 1964))

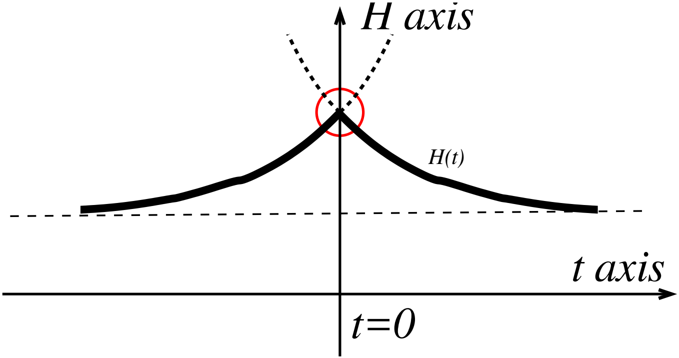
C. A distinguished applied mathematician's opinion

- " In fact, it requires no great mathematician to see that the reversibility theorem and Poincaré's recurrence theorem make irreversible behavior impossible for dynamical systems in the classical sense. Something must be added to the dynamics of conservative systems, something not consistent with it, in order to get irreversibility at all. That has long been clear to those competent in mechanics . . . ". (C Truesdell, p.375 of *Rational thermodynamics,* 2nd ed. (Springer 1984)).

A distinguished theoretical physicist's opinion

- He (Boltzmann) was forced to conclude that the irreversibility postulated by thermodynamics was incompatible with the reversible laws of dynamics (I. Prigogine, *Les Lois du Chaos* (Flammarion, Paris, 1994) cited by J Bricmont, *"Science of Chaos or Chaos in Science?"* Physicalia Magazine, 17, (1995) 3-4, pp.159-208.)

Initial conditions, probability



-
- Some of the (presumed) $H(t)$ curves for a gas which at the initial time $t = 0$ is required to be in a given macroscopic state (e.g. to have a given function $f(\cdot, \cdot, 0)$). If all the phase points compatible with that macroscopic state at time $t = 0$ are given equal probability-like weights (i.e. a uniform measure) then very nearly all the $H(t)$ functions will, it is believed, lie close to the heavy curve, which for $t > 0$ (only!) should agree with the actual behaviour.

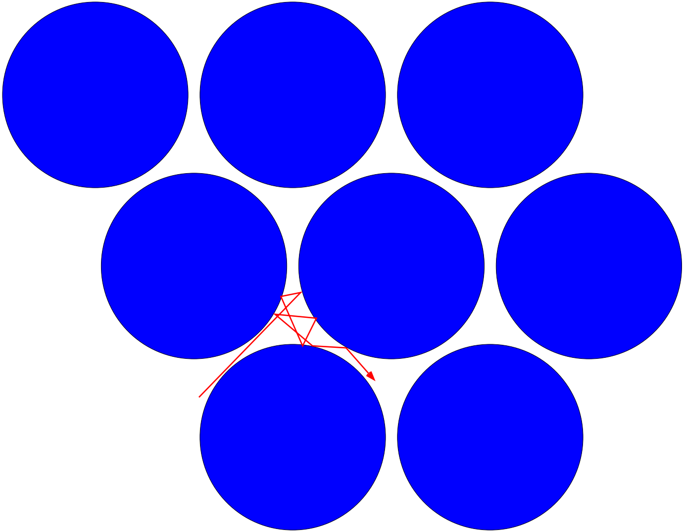
D. Explanation vs description

- example : Kepler and Newton
- Before trying to explain irreversibility, we need a good description.
- My proposal: characterize the "good" motions (i.e. the ones that could actually occur)

Two cases where it seems relatively easy to characterize "good" motions

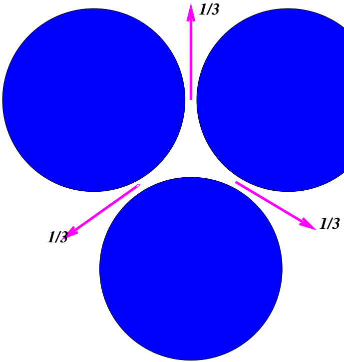
- A Lorentz gas with sufficiently densely packed scatterers (see following slides)
- A dilute gas of (e.g.) hard spheres.

Lorentz gas on a triangular lattice of hard disks



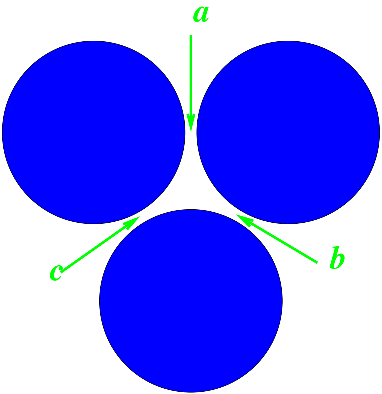
(Bunimovich and Sinai *Statistical properties of the Lorentz gas with periodic configuration of scatterers* Comm. Math. Phys. **78** 479-497 (1981))

output of a cell



the output fractions depend only on the geometry — not on what is happening outside the cell.

input of a cell



the input fractions $\frac{a}{a+b+c}$ etc do depend on what is happening outside the cell

What this example shows

- In a plausible forward motion of the Lorentz gas, the input ratios depend on what is going on outside (and hence, for example, on the time), but the output ratios are determined only by the geometry of the scatterers.
- In the reversed motion, it would be the the output ratios that depended on things besides the geometry of the scatterers
- In this model, the 'good' solutions of the equations of motion are distinguished by the fact that their output ratios are determined only by the geometry of the scatterers