

# Reversibility and irreversibility

Oliver Penrose  
Maxwell Institute and Department of Mathematics  
Heriot-Watt University  
Riccarton, Edinburgh EH14 4AS

August 4, 2009

*To appear in "PDE and Materials", report no. 44/2006 of the Mathematisches Forschungsinstitut Oberwolfach (ed. J.M.Ball, R.D.James and S.Müller)*

In order of increasing detail, some well-known forms of mathematical description for a material object or system of material objects are:

1. The thermodynamic description, which includes the principle that the entropy  $S$  of a thermally isolated system never decreases with time :

$$\text{if } t_1 < t_2 \text{ then } S(t_1) \leq S(t_2) \quad (1)$$

2. Macroscopic descriptions, such as the heat equation

$$\partial e / \partial t = K \nabla^2 T \quad (2)$$

3. Descriptions by kinetic equations such as Boltzmann's equation

$$\partial f / \partial t + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = I(f) \quad (3)$$

where  $I(f)$  is an integral expression representing the effect of collisions on the distribution function  $f(\mathbf{v}, \mathbf{x}, t)$  of particle velocities  $\mathbf{v}$  at position  $\mathbf{x}$ .

4. Microscopic representations as a system of particles obeying Newton's equations of motion

$$m d^2 \mathbf{x}_i / dt^2 = -\nabla_i V(\mathbf{x}_1, \dots, \mathbf{x}_N), \quad (4)$$

where  $V$  is the potential energy, or Schrödinger's equation for the same system

$$i\hbar \partial \psi / \partial t = \left\{ -(\hbar^2 / 2m) \sum \nabla_i^2 + V \right\} \psi \quad (5)$$

To verify consistency, one would like to be able to derive the less detailed models from the more detailed ones; for example it is easy to deduce (1) from (2). But in trying to derive any of the less detailed models (1,2,3) from microscopic

dynamics we encounter a difficulty: the microscopic equations (4) and (5) are ‘reversible’, i.e. invariant under the time reversal transformation

$$R: \quad t \rightarrow -t, \quad v \rightarrow -v, \quad d/dt \rightarrow -d/dt, \quad \psi \rightarrow \bar{\psi} \quad (6)$$

whereas the less detailed models are ‘irreversible’, as is the observed behaviour of real physical objects.

This paradox has generated much controversy and some strange remarks, even from scholars of high repute<sup>1</sup>. Mathematically, what the paradox tells us is this: anything that is derived from the microscopic equations (4) or (5) *alone* will have the same time-reversal symmetry as those equations. Since eqns (1,2,3) do not have that symmetry, it follows that they cannot be derived from microscopic mechanics *alone*. The model must contain, in addition, some non-mechanical ingredient<sup>2</sup>. What can this extra non-mechanical ingredient be?

The difficulty is that the equations of motion (4) and (5) have too many solutions. Some of these (sometimes called[2, 4] the ‘good’ solutions) correspond to plausible macroscopic behaviour, but others, the ‘bad’ solutions, correspond to macroscopic behaviour whose like we never see in the real world — heat flowing from cold to hot, for example. To complete the microscopic description we need some information to pick out the ‘good’ solutions and reject the ‘bad’ ones. This information will take the form of a plausible probability assumption.

To formulate the assumption, consider a thought-experiment. A container is separated into two equal parts by a partition. Initially (time  $t_0$ ) the left-hand compartment contains  $N$  molecules of gas and the right-hand one is empty. The partition is removed and the gas expands irreversibly into the empty compartment. Soon (time  $t_2$ ) it is distributed approximately evenly between the two compartments. Let us denote the number of molecules in the right-hand compartment by  $M$ . Initially the value of  $M$  is zero, but after the partition is removed we expect it to increase, eventually fluctuating around a mean value of  $\frac{1}{2}N$ . Correspondingly, the entropy would increase from 0 to a value fluctuating around  $k \log\{N!/(\frac{1}{2}N)!^2\} \approx Nk \log 2$ , in conformity with eqn (1).

To see how the ‘good’ solutions were picked out here, let us look at Bricmont’s account[2]: ‘the overwhelming majority of the microscopic configurations corresponding to the gas in the left half will evolve deterministically so as to induce the observed evolution of  $M$ . There may of course be some exceptional configurations, for which all the particles stay in the left half. All one is saying is that those configurations are extraordinarily rare, and that we do not expect to see even one of them appearing when we repeat the experiment many times, not even once “in a million years”, to put it mildly’.

In the above passage, starting from the fact that a certain special set of configurations (namely those which will evolve deterministically so as to give

---

<sup>1</sup>An example, from p. 375 of Truesdell[7]: ‘... 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.’

<sup>2</sup>“There cannot be a rigorous mechanical derivation of the macroscopic equations from the microscopic ones. Some additional information or assumption is indispensable” [5].

$M(t_2) = 0$ ) is in a tiny minority at time  $t_0$ , it is argued that the set of trajectories evolving from this special set of configurations has a tiny probability. This is tantamount to the following assumption: plausible probabilities for the evolution after time  $t_0$  can be obtained from a probability density in the space of dynamical states which at time  $t_0$  is uniform over the ‘initial’ set of dynamical states — those for which  $M = 0$ . Rather than pick out a specific class of ‘good’ solutions, the assumption proposes a probability measure under which (for large  $N$ ) ‘good’ solutions are highly probable. Notice the time-reversal asymmetry: the high-probability solutions are ‘good’ after time  $t_0$ , but not necessarily before.

Some of the conclusions that follow from such a probability assumption are:

- For finite systems, the principle of certain entropy increase (1) is replaced by a probabilistic statement<sup>3</sup>.
- Irreversibility is qualitative: the more particles in the system, the more irreversibly it behaves (i.e. the more unlikely is the reversed motion). ‘Absolute’ irreversibility can occur only in the limit of an infinite system.
- Any microscopic description using eqn (4) or (5) alone is incomplete: some probability information is needed as well, normally about the initial time<sup>4</sup>. For example, in Lanford’s derivation of Boltzmann’s equation (3) from microscopic dynamics[3], the particles are assumed to be as uncorrelated as possible at the initial time, and a similar independence assumption is made in Lebowitz and Spohn’s derivation of Fick’s law for self-diffusion[4].

*Note added after the talk.* A difficulty with the the probability assumption used above is that it refers to a special ‘initial’ time  $t_0$  with no statement of how this time is to be chosen or why it is so special<sup>5</sup>. Moreover, although the probability assumption implies (1) (in a probabilistic sense) in the special case where  $t_1 = t_0$ , it does not obviously do so for any other value of  $t_1$

This difficulty could be avoided by replacing the above probability assumption with the following stronger ‘multi-time’ assumption: given any time interval  $[t_1, t_2]$ , if a probability density in the space of dynamical states takes a simple form<sup>6</sup> at time  $t_1$ , then it induces a probability distribution on trajectories under which the trajectory is very likely to be ‘good’ on that time interval.

The multi-time probability assumption gives a family of microscopic probability densities for a given material object, each one giving a different mathematical model with its own ‘initial’ time  $t_1$  and its own time interval of validity.

<sup>3</sup>This was known to Boltzmann[1] :‘ My minimum theorem as well as the so-called Second Law of Thermodynamics are only theorems of probability’.

<sup>4</sup>Boltzmann again: ‘It can never be proved from the equations of motion alone that  $H$  [i.e.  $-S$ ] 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 haphazard [*sic*] governs freely, the probability that  $H$  decreases is always greater than that it increases.’ [1].

<sup>5</sup>This difficulty may be one of the reasons why some people (including the present author) have tried to trace irreversibility back to the truly special time when the world began. Whatever its merits, that enterprise does not help with the topic of this talk, which is the consistency of different mathematical models for the same material object.

<sup>6</sup>‘Simple’ needs a proper definition. It could mean, for example, ‘uniform over the part of the space compatible with the current macroscopic state’

These probability densities are not directly observable, and it is not claimed that any of them is the ‘true’ or ‘real’ probability density in dynamical space, whatever that means; all that is claimed is that the resulting mathematical models will all give good representations of the observed (probabilistic) behaviour over their respective time intervals of validity. That leaves us with a mathematical problem: to show that such a claim is consistent with the equations of mechanics (for a rudimentary discussion of this consistency problem, see [6].).

## References

- [1] L Boltzmann, On certain questions in the theory of gases, *Nature* **51**, 413-315(1895).
- [2] J Bricmont, Science of chaos or chaos in science?, *Physica Magazine* **17**, 3-4 pp 159-208 (1995) or <http://dogma.free.fr/txt/JB-Chaos.htm>
- [3] O E Lanford III, Time evolution of large classical systems, pp 1-111 of *Dynamical systems: theory and applications: Battelle Seattle Rencontres 1974*, ed. J. Moser (Springer Lecture Notes in Physics **38**, 1975).
- [4] J L Lebowitz and H Spohn, Microscopic basis for Fick’s law of self-diffusion, *J. Stat. Phys.* **28** 539-556 (1982); On the time evolution of macroscopic systems, *Comm. Pure. Appl. Math* **36**, 595-613 (1983).
- [5] N G van Kampen, p. 173 of *Fundamental problems in statistical mechanics*, ed. E.G.D.Cohen (North-Holland, Amsterdam 1962).
- [6] O Penrose, sec. IV.7 of *Foundations of statistical mechanics* (Dover 2005).
- [7] C Truesdell, p.375 of *Rational thermodynamics*, 2nd ed. (Springer 1984).