An asymmetric world

At the level of particles, things can happen in reverse, because particles obey time-symmetric laws of mechanics. But then why does matter, which is made up of these building blocks, behave irreversibly?

Oliver Penrose

My father once filmed a friend diving into a swimming pool. He liked to show the film in reverse, so that a big splash converged on the swimmer and expelled him from the water. The obvious impossibility of the reversed sequence of events illustrates the fact that our world is not symmetric under time reversal
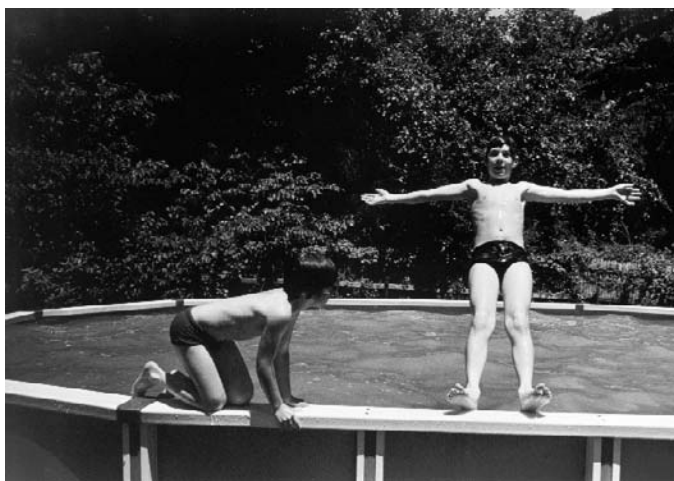
Most people are not worried by the asymmetry, but it does give physicists some trouble. Physicists want to explain the macroscopic behaviour of matter in terms of the microscopic mechanical laws obeyed by its component particles. These laws are symmetric under time reversal; for any motion obeying the laws of mechanics, the time-reversed motion also obeys these laws. Newton's laws have this symmetry, as do Maxwell's electromagnetic equations, Schrödinger's equation and Einstein's general relativity.

So, naively, one would expect the matter made out of the particles obeying these laws to behave in a time-symmetric way. Yet it does not. Divers are not expelled from swimming pools by converging splashes even though there is a dynamically possible motion of the molecules in the pool that would produce this effect. Nor do natural processes such as heat flow ever go into reverse. The question is, then, how to explain the irreversibility of the macroscopic behaviour of matter, given the reversibility of the microscopic laws obeyed by its constituent particles.

This 'paradox of irreversibility' came to light when James Clerk Maxwell and Ludwig Boltzmann formulated their kinetic theory of gases. Boltzmann's own response (in 1905) was: "From the fact that the differential equations of mechanics are left unchanged by reversing the sign of time ..., Herr Ostwald concludes that the mechanical view of the world cannot explain why natural processes run preferentially in a definite direction. But such a view appears to me to overlook that mechanical events are determined not only by differential equations but also by initial conditions."

Boltzmann's observation is certainly true, and many use it as the core of their explanations of irreversibility. But such an explanation seems to beg the question: why do we consider mechanical events to be determined by initial conditions, rather than by final conditions? Certainly, it is a fact of experience that you can start a simple mechanical system (a thrown ball, for example) in whatever state you choose, whereas the final state can be controlled, if at all, only by manipulating the initial state very accurately. But this special role of initial (as opposed to final) conditions in mechanics is another time-asymmetric feature of the world we live in — and would

Time's arrow: why can't you make a splash go backwards?

seem to require explanation along with the irreversibility of natural processes, rather than be used as part of that explanation.

The standard explanation of macroscopic irreversibility goes like this. Consider, as an example, a container separated into two equal-sized compartments by a partition. Initially one compartment contains gas and the other is empty. The partition is removed, and the gas expands irreversibly into the empty compartment. Soon it is distributed approximately evenly between the two compartments.

The explanation given is that the removal of the partition suddenly makes available to the system a much larger number of dynamical states (phase–space points). All but a tiny fraction of the dynamical states compatible with the initial macroscopic state (one compartment full, the other empty) lead to an approximately even distribution of gas between the two compartments, and therefore, so the argument goes, it is practically certain that the gas will end up approximately evenly distributed.

This argument is based on a probability assumption. The assumption is this: if practically all the possible initial dynamical states lead to a particular macroscopic behaviour, then it is practically certain that the actual system will exhibit that behaviour. The assumption is plausible enough, given our experience of such things as tossing coins. Only a tiny fraction of the ways of setting a spinning coin in motion will cause it to land on its edge, and our experience is that the coin is practically certain not to land on its edge.

But notice that use of future tense in "will exhibit". This signals another time-asymmetric feature of the world we live in, namely the time-asymmetric way that applied probability works. To my mind, this time asymmetry itself calls for explanation and should not lightly be incorporated into an explanation of the time asymmetry of something else.

The final stage in the standard discussion of the irreversibility paradox is to consider how the initial low-entropy state (a macroscopic state of tiny phase–space volume) was produced. It must have come from some even lower-entropy state, and so on back to the beginning of time, when the Universe was in a special macroscopic state of exceptionally low entropy. Some authors even seem to maintain that the low entropy of this primeval macroscopic state is, by itself, enough to account for the irreversibility of everything thereafter. But such a claim surely goes too far, since it takes for granted repeated applications of the asymmetric probability assumption just mentioned.

Moral: if you look at the world through rose-coloured spectacles, you cannot tell which parts of it really are rosy and which parts just look rosy. Likewise, if you use time-directed concepts like initial conditions and applied probability as tools for looking at the irreversibility paradox, you risk arriving at an incomplete or even misleading solution to this complex puzzle. ∎

Oliver Penrose is an emeritus professor in the Department of Mathematics at Heriot-Watt University, Edinburgh, and the author of *Foundations of Statistical Mechanics* (Dover, 2005).

ESSAY