



VINTAGE

CYCLES OF TIME

ROGER PENROSE

Contents

Cover

About the Book

About the Author

Also by Roger Penrose

Title Page

Preface

Prologue

Part 1. The Second Law and its underlying mystery

1.1 The relentless march of randomness

1.2 Entropy, as state counting

1.3 Phase space, and Boltzmann's definition of entropy

1.4 The robustness of the entropy concept

1.5 The inexorable increase of entropy into the future

1.6 Why is the past different?

Part 2. The oddly special nature of the Big Bang

2.1 Our expanding universe

2.2 The ubiquitous microwave background

2.3 Space-time, null cones, metrics, conformal geometry

2.4 Black holes and space-time singularities

2.5 Conformal diagrams and conformal boundaries

2.6 Understanding the way the Big Bang was special

Part 3. Conformal cyclic cosmology

3.1 Connecting with infinity

3.2 The structure of CCC

3.3 Earlier pre-Big-Bang proposals

3.4 Squaring the Second Law

3.5 CCC and quantum gravity

3.6 Observational implications

Epilogue

Notes

Acknowledgements

Appendices

Appendix A

Appendix B

Index

Copyright

About the Book

* Roger Penrose's groundbreaking and bestselling *The Road to Reality* provided a comprehensive yet readable guide to our present understanding of the laws that are currently believed to govern our universe. In *Cycles of Time*, he moves far beyond this to **develop a completely new perspective on cosmology, providing a quite unexpected answer to the often-asked question, 'What came before the Big Bang?'**

* The two key ideas underlying this novel proposal are **a penetrating analysis of the Second Law of thermodynamics** – according to which the 'randomness' of our world is continually increasing – and **a thorough examination of the *light-cone* geometry of space-time**. Penrose is able to combine these two central themes to show how the expected ultimate fate of our accelerating, expanding universe can actually be reinterpreted as the 'Big Bang' of a new one.

* On the way, many other basic ingredients are presented, and their roles discussed in detail, though without any complex mathematical formulae (these all being banished to the appendices). Various standard and non-standard cosmological models are presented, as is the fundamental and ubiquitous role of the cosmic microwave background. Also crucial to the discussion are the huge black holes lying in galactic centres, and their eventual disappearance via the mysterious process of Hawking evaporation.

About the Author

Professor Sir Roger Penrose is Emeritus Rouse Ball Professor of Mathematics at the University of Oxford. He has received a number of prizes and awards, including the 1988 Wolf Prize for physics which he shared with Stephen Hawking for their joint contribution to our understanding of the universe.

Also by Roger Penrose

*The Emperor's New Mind: Concerning Computers,
Minds, and the Laws of Physics*

*Shadows of the Mind: A Search for the Missing
Science of Consciousness*

*Road to Reality: A Complete Guide to the
Laws of the Universe*

ROGER PENROSE

Cycles of Time

An Extraordinary New View
of the Universe

VINTAGE BOOKS
London

Preface

ONE of the deepest mysteries of our universe is the puzzle of whence it came.

When I entered Cambridge University as a mathematics graduate student, in the early 1950s, a fascinating cosmological theory was in the ascendant, known as the steady-state model. According to this scheme, the universe had no beginning, and it remained more-or-less the same, overall, for all time. The steady-state universe was able to achieve this, despite its expansion, because the continual depletion of material arising from the universe's expansion is taken to be compensated by the continual creation of new material, in the form of an extremely diffuse hydrogen gas. My friend and mentor at Cambridge, the cosmologist Dennis Sciama, from whom I learnt the thrill of so much new physics, was at that time a strong proponent of steady-state cosmology, and he impressed upon me the beauty and power of that remarkable scheme of things.

Yet this theory has not stood the test of time. About 10 years after I had first entered Cambridge, and had become well acquainted with the theory, Arno Penzias and Robert Wilson discovered, to their own surprise, an all-pervading electromagnetic radiation, coming in from all directions, now referred to as the *cosmic microwave background* or CMB. This was soon identified, by Robert Dicke, as a predicted implication of the 'flash' of a *Big-Bang* origin to the universe, now presumed to have taken place some 14 thousand million years ago—an event that had been first seriously envisaged by Monsignor Georges Lemaître in 1927, as an implication of his work on Einstein's 1915

equations of general relativity and early observational indications of an expansion of the universe. With great courage and scientific honesty (when the CMB data became better established), Dennis Sciama publicly repudiated his earlier views and strongly supported the idea of the Big Bang origin to the universe from then on.

Since that time, cosmology has matured from a speculative pursuit into an exact science, and intense analysis of the CMB—coming from highly detailed data, generated by numerous superb experiments—has formed a major part of this revolution. However, many mysteries remain, and much speculation continues to be part of this endeavour. In this book, I provide descriptions not only of the main models of classical relativistic cosmology but also of various developments and puzzling issues that have arisen since then. Most particularly, there is a profound oddness underlying the Second Law of thermodynamics and the very nature of the Big Bang. In relation to this, I am putting forward a body of speculation of my own, which brings together many strands of different aspects of the universe we know.

My own unorthodox approach dates from the summer of 2005, though much of the detail is more recent. This account goes seriously into some of the geometry, but I have refrained from including, in the main body of the text, anything serious in the way of equations or other technicalities, all these being banished to the Appendices. The experts, only, are referred to those parts of the book. The scheme that I am now arguing for here is indeed unorthodox, yet it is based on geometrical and physical ideas which are very soundly based. Although something entirely different, this proposal turns out to have strong echoes of the old steady-state model!

I wonder what Dennis Sciama would have made of it.

Prologue

WITH his eyelids half closed, as the rain pelted down on him and the spray from the river stung his eyes, Tom peered into the swirling torrents as the water rushed down the mountainside. 'Wow', he said to his Aunt Priscilla, an astrophysics professor from the University of Cambridge, who had taken him to this wonderful old mill, preserved in excellent working order, 'is it always like this? No wonder all that old machinery can be kept buzzing around at such great speed.'

'I don't think it's always this energetic', said Priscilla, standing next to him behind the railing at the side of the river, and raising her voice somewhat, so as to be heard over the noise of the rushing water. 'The water's much more violent than usual, today, because of all this wet weather. You can see down there that a good portion of the water has had to be diverted away from the mill. Usually they would not do this, because they would have to make the most of a much more sedate flow. But now there's far more energy in the flow than is needed for the mill.'

Tom stared for some minutes into the wildly tumbling water and admired the patterns it made as it was flung into the air in sprays and convoluted surfaces. 'I can see there's a lot of power in that water, and I know that a couple of centuries ago the people were clever enough to see how all this energy could be used to drive these machines—doing the work of many human beings and making all that great woollen cloth. But where did the energy come from that got all that water high up on the mountain in the first place?'

‘The heat of the Sun caused the water in the oceans to evaporate and rise up into the air, so it would eventually come back down again in all this rain. So a good proportion of the rain would be deposited up high into the mountains’, replied Priscilla. ‘It’s really the energy from the Sun that is being harnessed to run the mill.’

Tom felt a little puzzled by this. He was often puzzled by the things that Priscilla told him, and was by nature often quite sceptical. He could not really see how just heat could lift water up into the air. And if there was all that heat around, why did he feel so cold now? ‘It was rather hot yesterday’, he grudgingly agreed. Though, still uneasy, he commented, ‘but I didn’t feel the Sun trying to lift me up into the air then, any more than I do now.’

Aunt Priscilla laughed. ‘No. it’s not really like that. It’s the tiny little molecules in the water in the oceans that the Sun’s heat causes to be more energetic. So these molecules then rush randomly around faster than they would otherwise, and a few of these “hot” molecules will move so fast that they break loose from the surface of the water and are flung into the air. And although there are only a relatively few molecules flung out at one time, the oceans are so vast that there would really be a lot of water flung up into the air altogether. These molecules go to make the clouds and eventually the water molecules fall down again as rain, a lot of which falls high in the mountains.’

Tom was still rather troubled, but at least the rain had now tapered off somewhat. ‘But this rain doesn’t feel at all hot to me.’

‘Think of the Sun’s heat energy first getting converted into the energy of rapid random motion of the water molecules. Then think of this rapid motion resulting in a small proportion of the molecules going so fast that they are flung high in the air in the form of water vapour. The energy of these molecules gets converted into what’s called gravitational potential energy. Think of throwing a ball up

into the air. The more energetically you throw it the higher it goes. But when it reaches its maximum height, it stops moving upwards. At that point its energy of motion has all been converted into this gravitational potential energy in its height above the ground. It's the same with the water molecules. Their energy of motion—the energy that they got from the Sun's heat—is converted into this gravitational potential energy, now at the top of the mountain, and when it runs down, this is converted back again into the energy in its motion, which is used to run the mill.'

'So the water isn't hot at all when it's up there?' asked Tom.

'Exactly, my dear. By the time that these molecules get very high in the sky, they slow down and often actually get frozen into tiny ice crystals—that's what most clouds are made of—so the energy goes into their height above the ground rather than into their heat motion. Accordingly, the rain won't be hot at all up there, and it's still quite cold even when it finally works its way down again, slowed down by the resistance of the air.'

'That's amazing!'

'Yes, indeed', and encouraged by the boy's interest, Aunt Priscilla eagerly took advantage of the opportunity to say more. 'You know, it's a curious fact that even in the *cold* water in this river there is still much more *heat* energy in the motion of the individual molecules running around randomly at great speed than there is in the swirling currents of water rushing down the mountainside!'

'Goodness. I'm supposed to believe that, am I?'

Tom thought for a few minutes, somewhat confused at first, but then rather attracted by what Priscilla said, remarked excitedly: 'Now you've given me a great idea! Why don't we build a special kind of mill that just directly uses all that energy of the motion of water molecules in some ordinary lake? It could use lots of tiny little windmill things, maybe like those things that spin in the wind, with

little cups on the ends so that they twirl round in the wind no matter which direction the wind is coming from. Only they'd be very tiny and in the water, so that the speed of the water molecules would spin them around, and you could use these to convert the energy in the motion in the water molecules to drive all sorts of machinery.'

'What a wonderful idea, Tom darling, only unfortunately it wouldn't work! That's because of a fundamental physical principle known as the Second Law of thermodynamics, which more or less says that things just get more and more disorganized as time goes on. More to the point, it tells you that you can't get useful energy out of the *random* motions of a hot—or cold—body, just like that. I'm afraid what you're suggesting is what they call a "Maxwell's demon".'

'Don't *you* start doing that! You know that Grandpa always used to call me a "little demon" whenever I had a good idea, and I didn't like it. And, that Second Law thing's not a very nice kind of law', Tom complained grumpily. Then his natural scepticism returned: 'And I'm not sure I can really believe in it anyway.' Then he continued 'I think laws like that just need clever ideas to get around them. In any case, I thought you said that it's the heat of the Sun that's responsible for heating the oceans and that it's that *random* energy of motion that flings it to the top of the mountain, and that's what's running the mill.'

'Yes, you're right. So the Second Law tells us that actually the heat of the Sun all by itself wouldn't work. In order to work, we *also* need the colder upper atmosphere, so that the water vapour can condense up above the mountain. In fact, the Earth as a whole doesn't get energy from the Sun overall.'

Tom looked at his aunt with a quizzical expression. 'What does the cold upper atmosphere have to do with it? Doesn't "cold" mean not so much energy as "hot"? How does a bit of "not-so-much energy" help? I don't get what you are saying at all. Anyway, I think you are contradicting

yourself', said Tom, gaining confidence in himself. 'First you tell me that the Sun's energy runs the mill, and now you tell me that the Sun doesn't give energy to the Earth after all!'

'Well, it doesn't. If it did, then the Earth would just keep on getting hotter and hotter as it gained energy. The energy that the Earth gets from the Sun in the daytime has all to go back into space eventually, which it does because of the cold night sky—except, I suppose, that with global warming, a little part of it does get held back by the Earth. It's because the Sun is a very hot spot in an otherwise cold dark sky ...'

Tom began to lose the thread of what she was saying and his mind began to wander. But he heard her say, '... so it's the manifest *organization* in the Sun's energy that enables us to keep the Second Law at bay.'

Tom looked at Aunt Priscilla, almost totally bemused. 'I don't think I really understand all that,' he said, 'and I don't see why I need to believe that "Second Law" thing in any case. Anyway, where does all that organization in the Sun come from? Your Second Law should be telling us that the Sun's getting more disorganized as time goes on, so it would have to have been enormously organized when it was first formed, since all the time it's sending out organization. Your "Second Law" thing tells us that its organization keeps getting lost.'

'It has to do with the Sun being such a hot spot in a dark sky. This extreme temperature imbalance provided the needed organization.'

Tom stared at Aunt Priscilla, with little comprehension, and now not really properly believing anything she was telling him. 'You tell me that counts as organization; well, I don't see why it should. All right, let's pretend it somehow does—but then you still haven't told me where that funny kind of organization comes from.'

‘From the fact that the gas that the Sun condensed from was previously spread uniformly, so that gravity could cause it to form clumps which condensed gravitationally into stars. A very long time ago, the Sun did just this; it condensed from this initially spread-out gas, getting hotter and hotter in the process.’

‘You’ll keep telling me one thing after another, going way back in time, but where does this thing you call “organization”, whatever it is, *originally* come from?’

‘Ultimately it comes from the Big Bang, which was what started the whole universe off with an utterly stupendous explosion.’

‘A thing like a big wallop explosion doesn’t sound like something organized. I don’t get it at all.’

‘You aren’t the only one! You’re in good company not to get it. Nobody *really* gets it. It’s one of the biggest puzzles of cosmology where the organization comes from, and in what way the Big Bang really represents organization in any case.’

‘Maybe there was something *more* organized *before* the Big Bang? That might do it.’

‘People have actually tried suggesting things like that for some while. There are theories in which our presently expanding universe had a previous collapsing phase which “bounced” to become our Big Bang. And there are other theories where little bits of a previous phase of the universe collapsed into things we call black holes, and these bits “bounced”, to become the seeds of lots and lots of new expanding universes, and there are others where new universes sprang out of things called “false vacuums”...’

‘That all sounds pretty crazy to me,’ Tom said.

‘And, oh yes, there’s another theory that I heard about recently ...’

Part 1

The Second Law and its underlying mystery

[1.1 The relentless march of randomness](#)

[1.2 Entropy, as state counting](#)

[1.3 Phase space, and Boltzmann's definition of entropy](#)

[1.4 The robustness of the entropy concept](#)

[1.5 The inexorable increase of entropy into the future](#)

[1.6 Why is the past different?](#)

1.1 The relentless march of randomness

THE SECOND LAW of thermodynamics—what law is this? What is its central role in physical behaviour? And in what way does it present us with a genuinely deep mystery? In the later sections of this book, we shall try to understand the puzzling nature of this mystery and why we may be driven to extraordinary lengths in order to resolve it. This will lead us into unexplored areas of cosmology, and to issues which I believe may be resolved only by a very radical new perspective on the history of our universe. But these are matters that will be our concern later. For the moment let us restrict our attention to the task of coming to terms with what is involved in this ubiquitous law.

Usually when we think of a ‘law of physics’ we think of some assertion of equality between two different things. Newton’s second law of motion, for example, equates the rate of change of momentum of a particle (momentum being mass times velocity) with the total force acting upon it. As another example, the law of conservation of energy asserts that the total energy of an isolated system at one time is equal to its total energy at any other time. Likewise, the law of conservation of electric charge, of momentum, and of angular momentum, each asserts a corresponding equality for the total electric charge, for the total momentum, and for total angular momentum. Einstein’s famous law $E=mc^2$ asserts that the energy of a system is always equal to its mass multiplied by the square of the speed of light. As yet another example, Newton’s third law

asserts that the force exerted by a body A on a body B, at any one time, is always equal and opposite to the force acting *on* A due to B. And so it is for many of the other laws of physics.

These are all *equalities*—and this applies also to what is called the *First* Law of thermodynamics, which is really just the law of conservation of energy again, but now in a thermodynamic context. We say ‘thermodynamic’ because the energy of the *thermal motions* is now being taken into account, i.e. of the random motions of individual constituent particles. This energy is the *heat* energy of a system, and we define the system’s *temperature* to be this energy per degree of freedom (as we shall be considering again later). For example, when the friction of air resistance slows down a projectile, this does not violate the full conservation law of energy (i.e. the First Law of thermodynamics)—despite the loss of kinetic energy, due to the projectile’s slowing—because the air molecules, and those in the projectile, become slightly more energetic in their random motions, from *heating* due to the friction.

However, the *Second* Law of thermodynamics is not an equality, but an *inequality*, asserting merely that a certain quantity referred to as the *entropy* of an isolated system—which is a measure of the system’s disorder, or ‘randomness’—is *greater* (or at least not smaller) at later times than it was at earlier times. Going along with this apparent weakness of statement, we shall find that there is also certain vagueness or subjectivity about the very definition of the entropy of a general system. Moreover, in most formulations, we are led to conclude that there are occasional or exceptional moments at which the entropy must be regarded as actually (though temporarily) *reducing* with time (in a fluctuation) despite the general trend being that the entropy increases.

Yet, set against this seeming imprecision inherent in the Second Law (as I shall henceforth abbreviate it), this law

has a universality that goes far beyond any particular system of dynamical rules that one might be concerned with. It applies equally well, for example, to relativity theory as it does to Newtonian theory, and also to the continuous fields of Maxwell's theory of electromagnetism (that we shall be coming to briefly in [§2.6](#), [§3.1](#) and [§3.2](#), and rather more explicitly in [Appendix A1](#)) just as well as it does to theories involving only discrete particles. It applies also to hypothetical dynamical theories that we have no good reason to believe have relevance to the actual universe that we inhabit, although it is most pertinent when applied to realistic dynamical schemes, such as Newtonian mechanics, which have a *deterministic* evolution and are *reversible in time*, so that for any allowed evolution into the future, reversing the time direction gives us another equally allowable evolution according to the dynamical scheme.

To put things in familiar terms, if we have a moving-picture film depicting some action that is in accordance with dynamical laws—such as Newton's—that are reversible in time, then the situation depicted when the film is run in reverse will also be in accordance with these dynamical laws. The reader might well be puzzled by this, for whereas a film depicting an egg rolling off a table, falling to the ground, and smashing would represent an allowable dynamical process, the time-reversed film—depicting the smashed egg, originally as a mess on the floor, miraculously assembling itself from the broken pieces of shell, with the yolk and albumen separately joining up to become surrounded by the self-assembling shell, and then jumping up on to the table—is not an occurrence that we expect ever to see in an actual physical process ([Fig. 1.1](#)). Yet the full Newtonian dynamics of each individual particle, with its accelerated response (in accordance with Newton's second law) to all forces acting upon it, and the elastic reactions involved in any collision between constituent

particles, is completely reversible in time. This also would be the case for the refined behaviour of relativistic and quantum-mechanical particles, according to the standard procedures of modern physics—although there are some subtleties arising from the black-hole physics of general relativity, and also with regard to quantum mechanics, that I do not wish to get embroiled in just yet. Some of these subtleties will actually be crucially important for us later, and will be considered particularly in [§3.4](#). But for the moment, an entirely Newtonian picture of things will suffice.

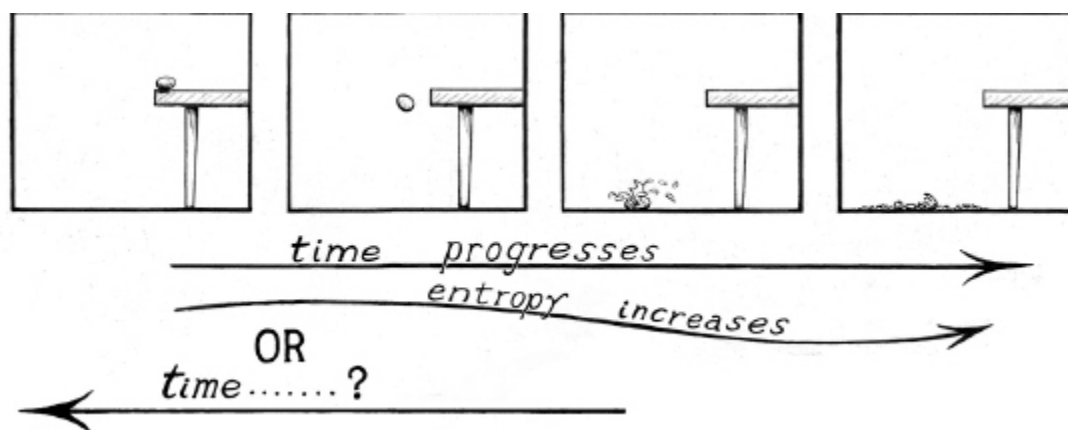


Fig. 1.1 An egg rolling off a table, falling to the ground and smashing according to time-reversible dynamical laws.

We have to accustom ourselves to the fact that the situations that are depicted by *both* directions of film-running are consistent with Newtonian dynamics, but the one showing the self-assembling egg depicts an occurrence that is inconsistent with the Second Law, and would be such an enormously improbable sequence of events that we can simply reject it as a realistic possibility. What the Second Law indeed states, roughly speaking, is that things are getting more ‘random’ all the time. So if we set up a particular situation, and then let the dynamics evolve it into the future, the system will evolve into a more random-looking state as time progresses. Strictly, we should not say

that it *will* evolve into a more random-looking state but that, in accordance with what has been said above, it is (something like) *overwhelmingly likely* to evolve into such a more random state. In practice, we must expect that, according to the Second Law, things are indeed getting progressively more and more random with time, but that this represents merely an overwhelming probability, not quite an absolute certainty.

Nevertheless we can assert, with a considerable amount of confidence, that what we shall experience will be an entropy increase—in other words an increase in randomness. Stated that way, the Second Law sounds perhaps like a council of despair, for it tells us that things are just getting more and more disorganized as time progresses. This does not sound like any kind of a mystery, however, as the title of [Part 1](#) seems to be suggesting that it should. It's just an obvious feature of the way things would behave if left entirely to themselves. The Second Law appears to be just expressing an inevitable and perhaps depressing feature of everyday existence. Indeed, from this point of view, the Second Law of thermodynamics is one of the most natural things imaginable, and certainly something that reflects a completely commonplace experience.

Some might worry that the emergence of life on this Earth, with its seemingly unbelievable sophistication, represents a contradiction with this increase of disorder that the Second Law demands. I shall be explaining later ([see §2.2](#)) why there is in fact no contradiction. Biology is, as far as we know, entirely consistent with the overall entropy increase that the Second Law demands. The mystery referred to in the title of [Part 1](#) is a mystery of *physics* of an entirely different order of scale. Although it has some definite relation to that mysterious and puzzling organization that we are continually being presented with

through biology, we have good reason to expect that the latter presents no paradox with regard to the Second Law.

One thing should be made clear, however, with regard to the Second Law's physical status: it represents a separate principle that must be *adjoined* to the dynamical laws (e.g. to Newton's laws), and is not to be regarded as a *deduction* from them. The actual *definition* of the entropy of a system *at any one moment* is, however, symmetrical with regard to the direction of time (so we get the same entropy definition, for our filmed falling egg, at any one moment, irrespective of the direction in which the film is shown), and if the dynamical laws are also symmetrical in time (as is indeed the case with Newtonian dynamics), the entropy of a system being not always constant in time (as is clearly so with the smashing egg), then the Second Law cannot be a deduction from these dynamical laws. For if the entropy is increasing in a particular situation (e.g. egg smashing), this being in accordance with the Second Law, then the entropy must be *decreasing* in the reversed situation (egg miraculously assembling), which is in gross violation of the Second Law. Since both processes are nevertheless consistent with the (Newtonian) dynamics, we conclude that the Second Law cannot simply be a *consequence* of the dynamical laws.

1.2 Entropy, as state counting

BUT HOW DOES the physicist's notion of 'entropy', as it appears in the Second Law, actually *quantify* this 'randomness', so that the self-assembling egg can indeed be seen to be overwhelmingly improbable, and thereby rejected as a serious possibility? In order to be a bit more explicit about what the entropy concept actually is, so that we can make a better description of what the Second Law actually asserts, let us consider a physically rather simpler example than the breaking egg. The Second Law tells us, for example, that if we pour some red paint into a pot and then some blue paint into the same pot and give the mixture a good stir, then after a short period of such stirring the different regions of red and of blue will lose their individuality, and ultimately the entire contents of the pot will appear to have the colour of a uniform purple. It seems that no amount of further stirring will convert the purple colour back to the original separated regions of red and blue, despite the time-reversibility of the submicroscopic physical processes underlying the mixing. Indeed, the purple colour should eventually come about spontaneously, even without the stirring, especially if we were to warm the paint up a little. But with stirring, the purple state is reached much more quickly. In terms of entropy, we find that the original state, in which there are distinctly separated regions of red and blue paint, will have a relatively low entropy, but that the pot of entirely purple paint that we end up with will have a considerably larger entropy. Indeed, the whole stirring procedure provides us with a situation that is not only consistent with the Second

Law, but which begins to give us a feeling of what the Second Law is all about.

Let us try to be more precise about the entropy concept, so that we can be more explicit about what is happening here. What actually *is* the entropy of a system? Basically, the notion is a fairly elementary one, although involving some distinctly subtle insights, due mainly to the great Austrian physicist Ludwig Boltzmann, and it has to do just with *counting* the different possibilities. To make things simple, let us idealize our pot of paint example so that there is just a (very large) finite number of different possibilities for the locations of each molecule of red paint or of blue paint. Let us think of these molecules as red balls or blue balls, these being allowed to occupy only discrete positions, centred within N^3 cubical compartments, where we are thinking of our paint pot as an enormously subdivided $N \times N \times N$ cubical crate composed of these compartments ([see Fig. 1.2](#)), where I am assuming that every compartment is occupied by exactly one ball, either red or blue (represented as white and black, respectively, in the figure).

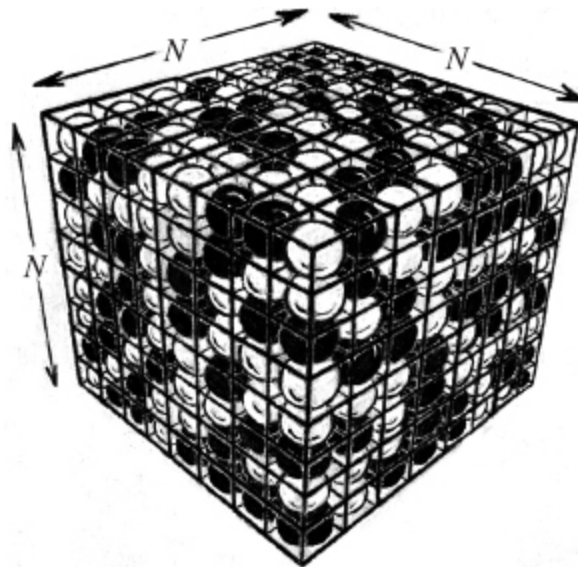


Fig. 1.2 $N \times N \times N$ cubical crate, each compartment containing a red or blue ball.

To judge the colour of the paint at some place in the pot, we make some sort of average of the relative density of red balls to blue balls in the neighbourhood of the location under consideration. Let us do this by containing that location within a cubical box that is much smaller than the entire crate, yet very large as compared with the individual cubical compartments just considered. I shall suppose that this box contains a large number of the compartments just considered, and belongs to a cubical array of such boxes, filling the whole crate in a way that is less refined than that of the original compartments ([Fig. 1.3](#)). Let us suppose that each box has a side length that is n times as great as that of the original compartments, so that there are $n \times n \times n = n^3$ compartments in each box. Here n , though still very large, is to be taken to be far smaller than N :

$$N \gg n \gg 1.$$

To keep things neat, I suppose that N is an exact multiple of n , so that

$$N = kn$$

where k is a whole number, giving the number of boxes that span the crate along each side. There will now be $k \times k \times k = k^3$ of these intermediate-sized boxes in the entire crate.

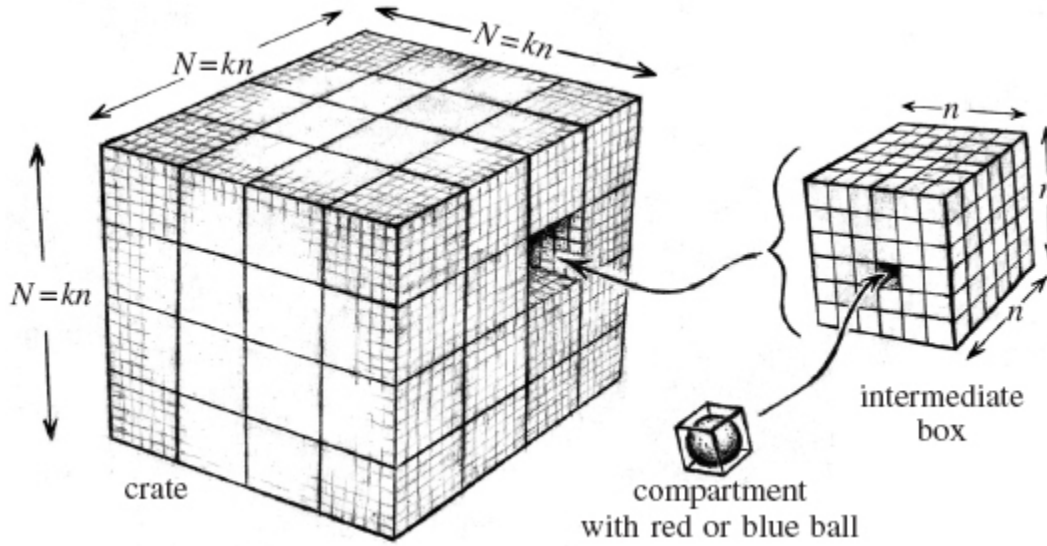


Fig. 1.3 The compartments are grouped together into k^3 boxes, each of size $n \times n \times n$.

The idea will be to use these intermediate boxes to provide us with a measure of the 'colour' that we see at the location of that box, where the balls themselves are considered to be too small to be seen individually. There will be an average colour, or *hue* that can be assigned to each box, given by 'averaging' the colours of the red and blue balls within that box. Thus, if r is the number of red balls in the box under consideration, and b the number of blue balls in it (so $r+b = n^3$), then the hue at that location is taken to be defined by the ratio of r to b . Accordingly, we consider that we get a redder hue if r/b is larger than 1 and a bluer hue if r/b is smaller than 1.

Let us suppose that the mixture *looks* to be a uniform purple to us if every one of these boxes of $n \times n \times n$ compartments has a value of r/b that is between 0.999 and 1.001 (so that r and b are the same, to an accuracy of one tenth of a per cent). This may seem, at first consideration, to be a rather stringent requirement (having to apply to *every* individual $n \times n \times n$ compartment). But when n gets very large, we find that the vast majority of the ball arrangements *do* satisfy this condition! We should also bear

in mind that when considering molecules in a can of paint, the number of them will be staggeringly large, by ordinary standards. For example, there could well be something like 10^{24} molecules in an ordinary can of paint, so taking $N=10^8$ would not be at all unreasonable. Also, as will be clear when we consider that colours look perfectly good in digitally displayed photographs with a pixel size of only 10^{-2} cm, taking a value of $k=10^3$ is also very reasonable, in this model. From this, we find that, with these numbers ($N=10^8$ and $k=10^3$, so $n=10^5$) there are around 10^{23} 570000 000000 000000 different arrangements of the entire collection of $\frac{1}{2}N^3$ red balls and $\frac{1}{2}N^3$ blue balls that give the appearance of a uniform purple. There are only a mere 10^{46} 500000 000000 different arrangements which give the original configuration in which the blue is entirely at the top and the red entirely at the bottom. Thus, for balls distributed entirely at random, the probability of finding uniform purple is a virtual certainty, whereas the probability of finding all the blue ones at the top is something like 10^{-23} 570000 000000 000000 (and this figure is not substantially changed if we do not require ‘all’ the blue balls to be initially at the top but, say, only 99.9% of them to be at the top).

We are to think of the ‘entropy’ to be something like a measure of these probabilities or, rather, of these different numbers of arrangements that give the same ‘overall appearance’. Actually, to use these numbers directly would give an exceedingly unruly measure, owing to their vast differences in size. It is fortunate, therefore, that there are good theoretical reasons for taking the (natural) *logarithm* of these numbers as a more appropriate ‘entropy’ measure. For those readers who are not very familiar with the notion of a logarithm (especially a ‘natural’ logarithm), let us phrase things in terms of the logarithm taken to the base 10—referred to here as ‘log₁₀’ (rather than the natural

logarithm, used later, which I refer to simply as 'log'). To understand \log_{10} , the basic thing to remember is that

$$\log_{10} 1=0, \log_{10} 10=1, \log_{10} 100=2, \log_{10} 1000=3, \\ \log_{10} 10000=4, \text{ etc.}$$

That is, to obtain the \log_{10} of a power of 10, we simply count the number of 0s. For a (positive) whole number that is not a power of 10, we can generalize this to say that the integral part (i.e. the number before the decimal point) of its \log_{10} is obtained by counting the total number of digits and subtracting 1, e.g. (with the integral part printed in bold type)

$$\log_{10} 2=\mathbf{0}.301\,029\,995\,66\ldots \\ \log_{10} 53=\mathbf{1}.724\,275\,869\,60\ldots \\ \log_{10} 9140=\mathbf{3}.960\,946\,195\,73\ldots$$

etc., so in each case the number in bold type is just one less than the number of digits in the number whose \log_{10} is being taken. The most important property of \log_{10} (or of log) is that it *converts multiplication to addition*; that is:

$$\log_{10} (ab) = \log_{10} a + \log_{10} b.$$

(In the case when a and b are both powers of 10, this is obvious from the above, since multiplying $a=10^A$ by $b=10^B$ gives us $ab=10^{A+B}$.)

The significance of the above displayed relation to the use of the logarithm in the notion of entropy is that we want the entropy of a system which consists of two separate and completely independent components to be what we get by simply *adding* the entropies of the individual parts. We say that, in this sense, the entropy

concept is *additive*. Indeed, if the first component can come about in P different ways and the second component in Q different ways, then there will be the product PQ of different ways in which the entire system—consisting of both components together—can come about (since to each of the P arrangements giving the first component there will be exactly Q arrangements giving the second). Thus, by defining the entropy of the state of any system to be proportional to the *logarithm* of the number of different ways that that state can come about, we ensure that this additivity property, for independent systems, will indeed be satisfied.

I have, however, been a bit vague, as yet, about what I mean by this ‘number of ways in which the state of a system can come about’. In the first place, when we model the locations of molecules (in a can of paint, say), we would normally not consider it realistic to have discrete compartments, since in Newtonian theory there would, in full detail, be an infinite number of different possible locations for each molecule rather than just a finite number. In addition, each individual molecule might be of some asymmetrical shape, so that it could be oriented in space in different ways. Or it might have other kinds of internal degrees of freedom, such as distortions of its shape, which would have to be correspondingly taken into account. Each such orientation or distortion would have to count as a different configuration of the system. We can deal with all these points by considering what is known as the *configuration* space of a system, which I next describe.

For a system of d degrees of freedom, the configuration space would be a d -dimensional space. For example, if the system consisted of q point particles p_1, p_2, \dots, p_q (each without any internal degrees of freedom), then the configuration space would have $3q$ dimensions. This is because each individual particle requires just three

coordinates to determine its position, so there are $3q$ coordinates overall, whereby a single point P of configuration space defines the locations of all of p_1, p_2, \dots, p_q together ([see Fig. 1.4](#)). In more complicated situations, where there are internal degrees of freedom as above, we would have more degrees of freedom for each such particle, but the general idea is the same. Of course, I am not expecting the reader to be able to ‘visualize’ what is going on in a space of such a high number of dimensions. This will not be necessary. We shall get a good enough idea if we just imagine things going on in a space of just 2 dimensions (such as a region drawn on a piece of paper) or of some region in ordinary 3-dimensional space, provided that we always bear in mind that such visualizations will inevitably be limited in certain ways, some of which we shall be coming to shortly. And of course we should always keep in mind that such spaces are purely abstract mathematical ones which should not be confused with the 3-dimensional *physical* space or 4-dimensional *physical* space-time of our ordinary experiences.