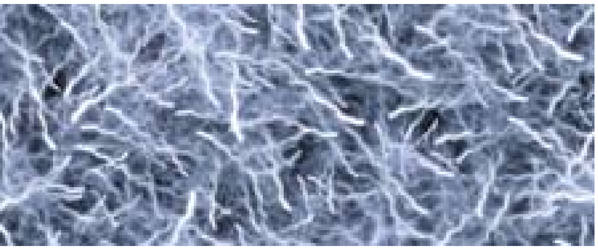


▲ Cosmic chaos
The chaotic, almost infinitely small Universe, a tiny fraction of a second after the Big Bang. In this representation the bright lines represent shortlived particles that are continually created and annihalated.

▶▶ Andromeda Galaxy (M31)
Looking out through the red foreground stars of our Galaxy, we see our nearest large neighbour in its entirety, comprising perhaps 200 billion stars. In the ultraviolet vision of the GALEX satellite, star-forming regions are highlighted in blue, whereas older parts of the system, which have converted their gas into stars, are yellow or orange. At a distance of 2.2 million light-years, it is the most distant object visible to the naked human eye.

Everything – space, time, matter – came into existence with a 'Big Bang', around 13.7 thousand million years ago. The Universe then was a strange place, as alien as it could possibly be. There were no planets, stars or galaxies; there was only a melée of elementary particles; these filled the Universe. Moreover, the entire Universe was smaller than a pin-prick, and it was unbelievably hot. At once it began to expand – and as it spread out and cooled from this bizarre, unexpected start, it evolved into the Universe we see today.

Modern science is unable to describe or explain anything that happened in the first 10^{-43} seconds after the Big Bang. This interval, 10^{-43} seconds, is known as the Planck time, named after the German scientist Max Karl Ernst Planck. He was the first to introduce the concept of energy not as a continuous flow, but in packets, or 'quanta', each with a specific energy. Quantum theory is now at the base of much of modern physics; it deals with the Universe on the smallest scales and is certainly one of the two great achievements of twentieth century theoretical science. The other is Einstein's general theory of relativity, which deals with the physics of very large scales – astronomical scales, in fact. Despite the fact that, in their own realms, both theories are extremely well tested by experiments and observations, there are major problems in reconciling these theories with each other. In particular, they treat time in fundamentally different ways. Einstein's theories treat time as a coordinate; it is therefore continuous, and we move smoothly from one moment to the next. In quantum theory, however, the Planck time represents a fundamental limit – the smallest unit of time that can be said to have any meaning at all, and the smallest unit that could ever, even in theory, be measured. Even if we built the most accurate clock possible, we would see it jump rather erratically from one Planck time to the next.

Attempting to reconcile these two contrasting views of time is one of the major challenges for 21st century physics. For now, in the small, hot, dense Universe that existed just after the Big Bang, quantum physics holds sway, and hence we begin our scientific study of the Universe 10^{-43} seconds after the beginning.

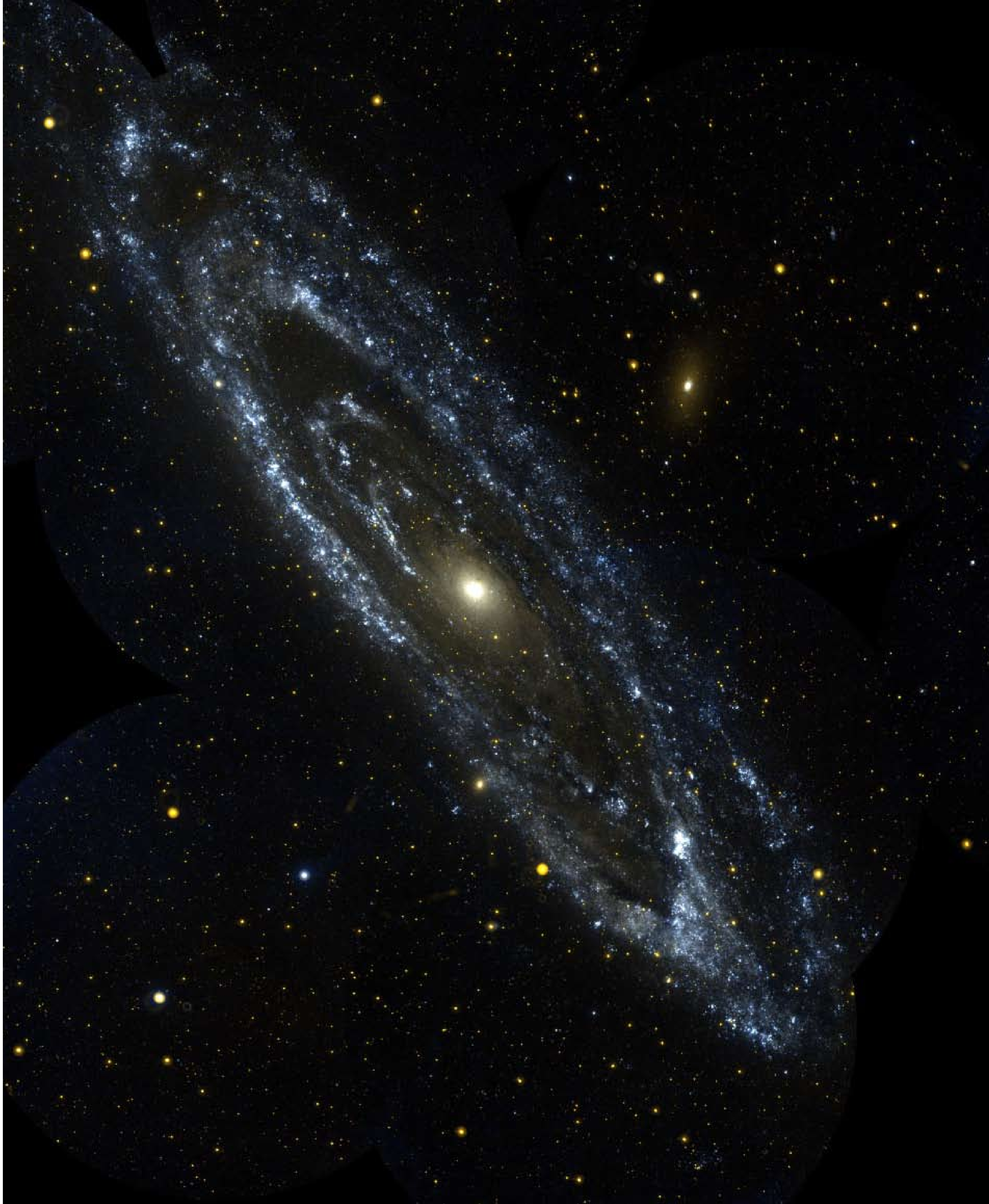
The Big Bang is a counterintuitive idea. Our common sense seems much more impressed by a static and infinite Universe, and yet there are good scientific reasons to believe in this singular event. If we accept the Big Bang, it is possible to trace the whole sequence of events from that first Planck time right up until the present day, where we find ourselves on what Carl Sagan memorably described as our 'pale blue dot'.

The beginning of time

So let us look back to the very start of the Universe – just after the Big Bang itself. It is tempting to picture the Universe suddenly bursting out in a vast ocean of space, but this is completely misleading. The true picture of the Big Bang is one in which space, matter and, crucially, time were born and started together. Space did not appear out of 'nothingness';

Standard form

10^{-43} is a convenient form of mathematical shorthand, and could be written as a decimal point followed by 42 zeros and then a 1. This is precise enough, but is decidedly clumsy (0.000000 000000000000000000000000000000000001) In order to deal with the incredibly large and incredibly small numbers that pervade astronomy, we will use so-called standard form throughout the book. 10^{33}, for example, is 1 followed by 33 zeros – we could write 1000 as 10^3, or 0.001 as 10^{-2}.



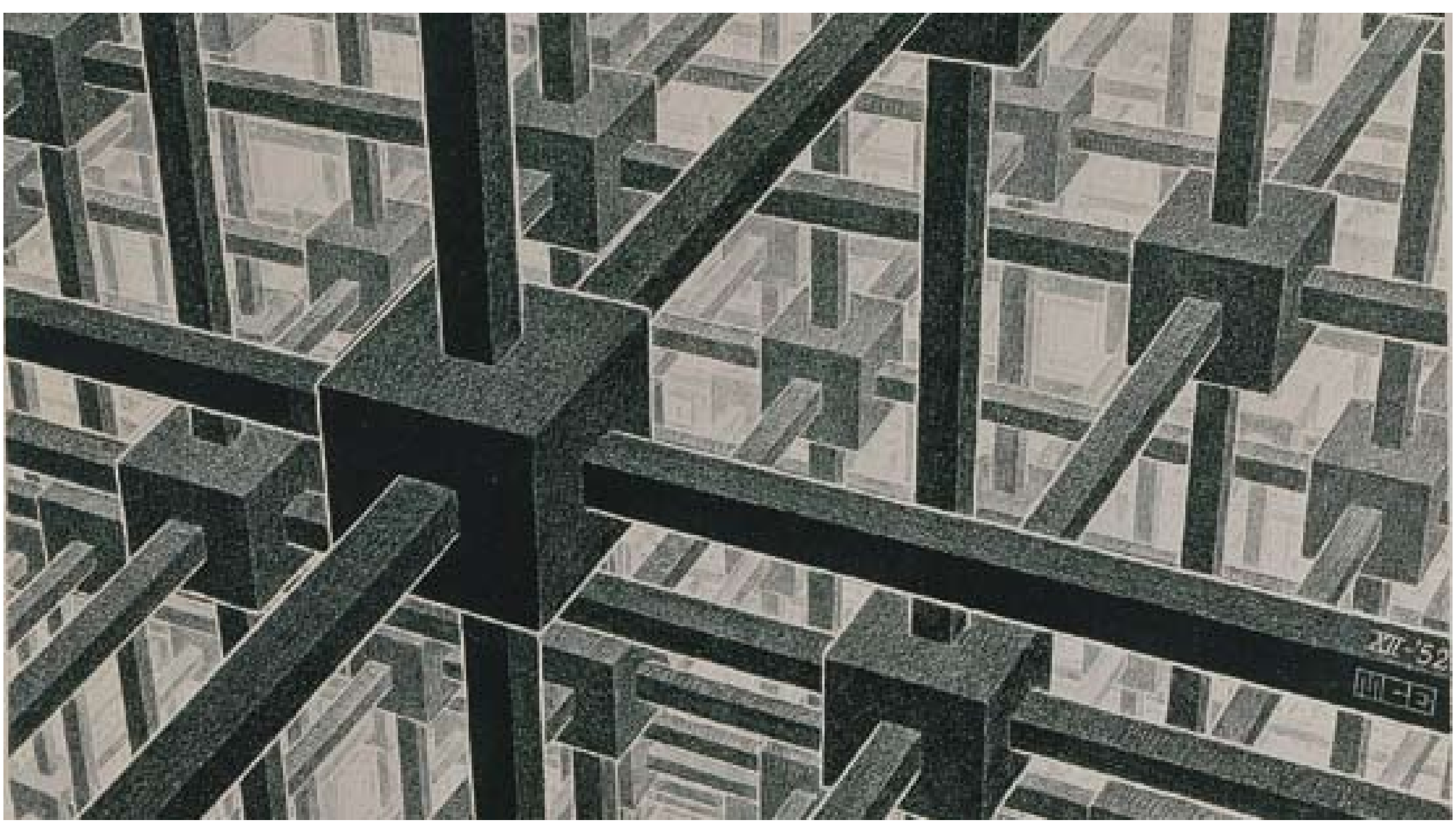
▶▶ Large-scale structure in the Universe
This enormous cluster of galaxies (AC03627), 250 million light-years away, is typical of what we can see in any direction provided we avoid the dust and gas of our own Galaxy and its neighbours. Clusters such as these are the largest objects to be held together by their own gravitational pull.

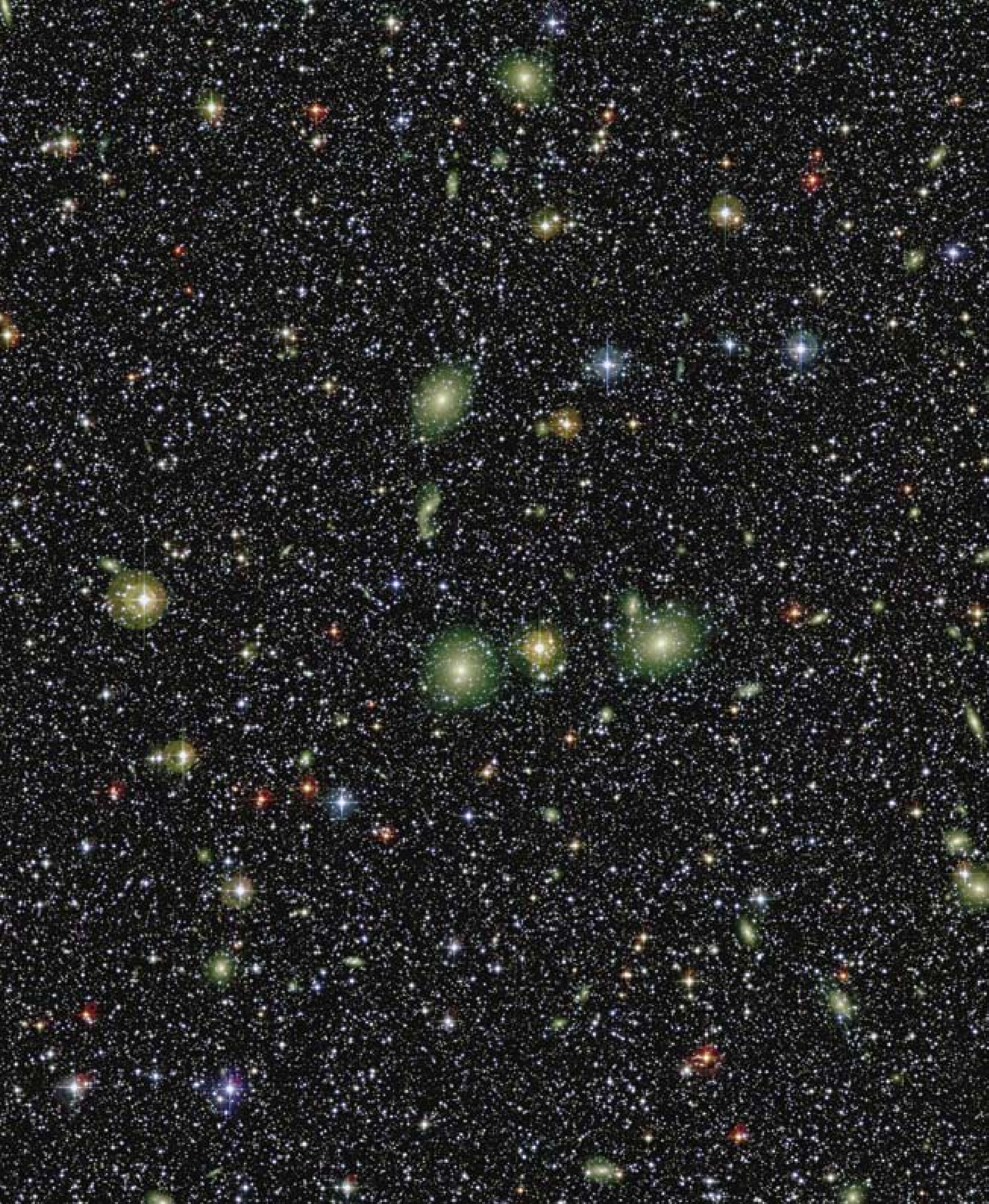
before the moment of creation there was no 'nothingness'. Time itself had not yet begun, and so it does not even make sense to speak of a time before the Big Bang. Not even a Shakespeare or an Einstein could explain this in plain English, though the combination of the two might be useful!

It also follows that when we survey the Universe today, it is meaningless to ask just 'where' the Big Bang happened. Space only came into existence with the Big Bang itself. Hence, in those first few fractions of a second, the entire Universe we see today was in a tiny region, smaller than an atomic nucleus; the Big Bang happened 'everywhere', and there was no central point. A nice illustration of this is given in a famous painting by Escher, known somewhat unromantically as 'Cubic Space Division'. Imagine standing on any of the cubes which mark junctions on this lattice, while each and every one of the rods joining the cubes expands. From your perspective, it would seem that everything is rushing away from you, and it might seem natural at first to conclude that you are in a special location – the centre of the expansion. Yet standing back and thinking allows you to realise that the expansion would look the same wherever in the lattice you were; there is no centre. The situation is very similar in our Universe; each group of galaxies appears to be rushing away from us, and yet an observer looking back at us from these distant stars would see the same illusion and would presumably be just as likely to conclude that he is at the centre of the expansion.

Another problem concerns the often-asked, and at first glance sensible question 'How big is the Universe?' Here we again have a major problem – there seem to be two possible answers. Either the Universe is of finite size, or else it isn't. If finite, what lies outside it? The question is meaningless – space itself exists only within the Universe, and therefore

▼ Cubic Space Division
Dutch artist M. C. Escher created this lithograph in 1952. Born in Holland in 1898, his work became widely recognized after his first important exhibition was written up in *Time* magazine in 1956. Mathematicians recognized the extraordinary visualization of their principles in his work.

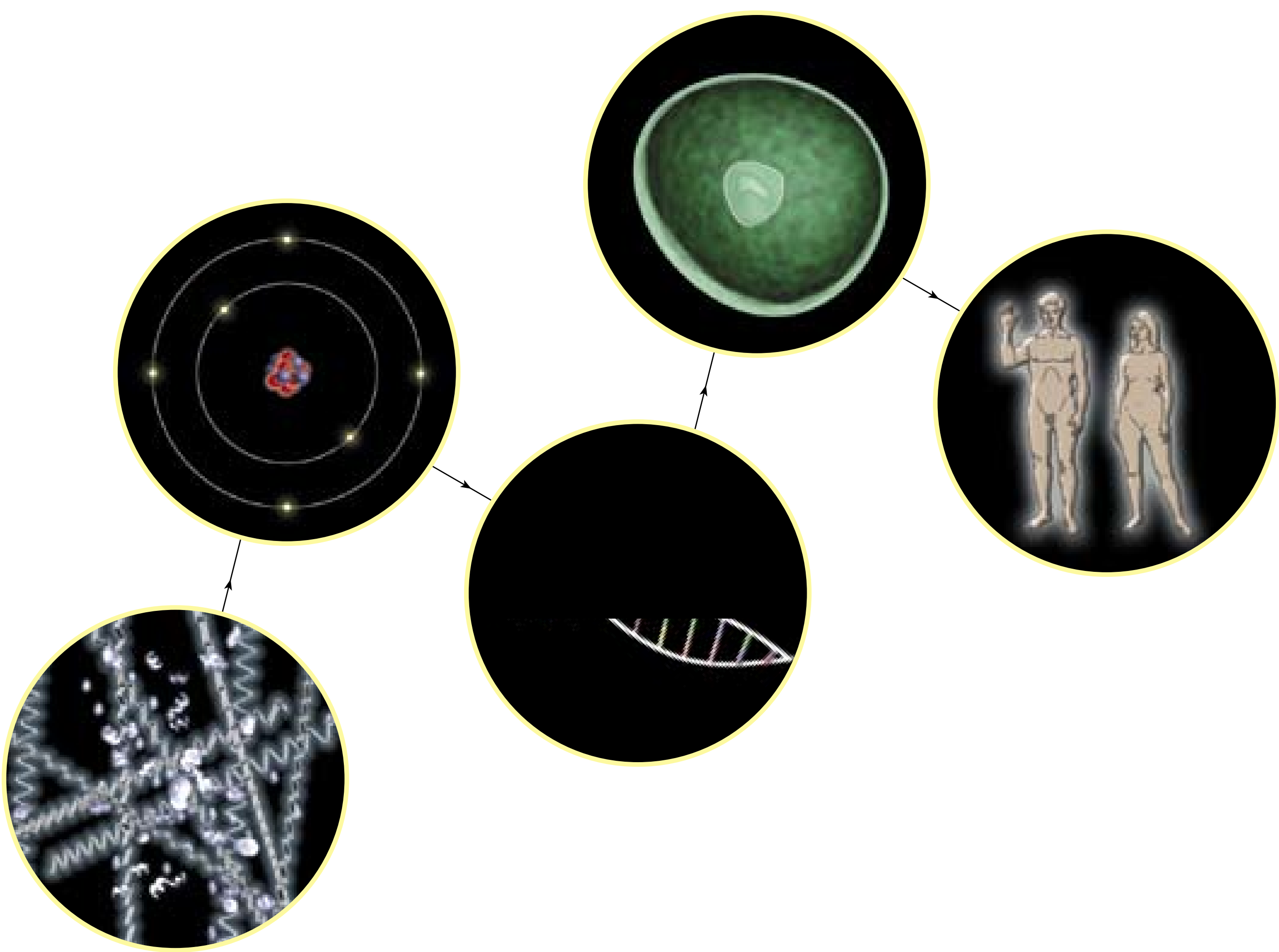




there is literally no 'outside'. On the other hand, to say the Universe is infinite is really to say that it size is not definable. We cannot explain infinity in everyday language, and neither could Einstein (we know, because Patrick asked him!).

Remember, too, that we need to consider time as a coordinate; in other words, we cannot simply ask 'how big is the Universe?' as the answer will change over time. We could ask 'how big is the Universe now?' but as we shall later, a consequence of relativity is that is impossible to define a single moment called 'now' which has the same meaning across the entire Universe.

Talking about a Universe which has a particular size immediately leads to thoughts of an edge. If we travelled far enough, would we hit a brick wall? The answer is 'no' - the Universe is what mathematicians call finite but unbounded. A useful analogy is that of an ant crawling on a ball. By travelling always in the same direction it will never come up against a barrier, and it can cover an infinite distance (provided that no-one interferes and kicks the ball away). This is despite the finite size of the ball, to which the ant will be completely oblivious. Similarly, if we were to set off in a powerful spacecraft in a straight line we would never reach the edge of the Universe – but this does not mean that the Universe is infinite. The analogy is limited, because it does not involve more than the two

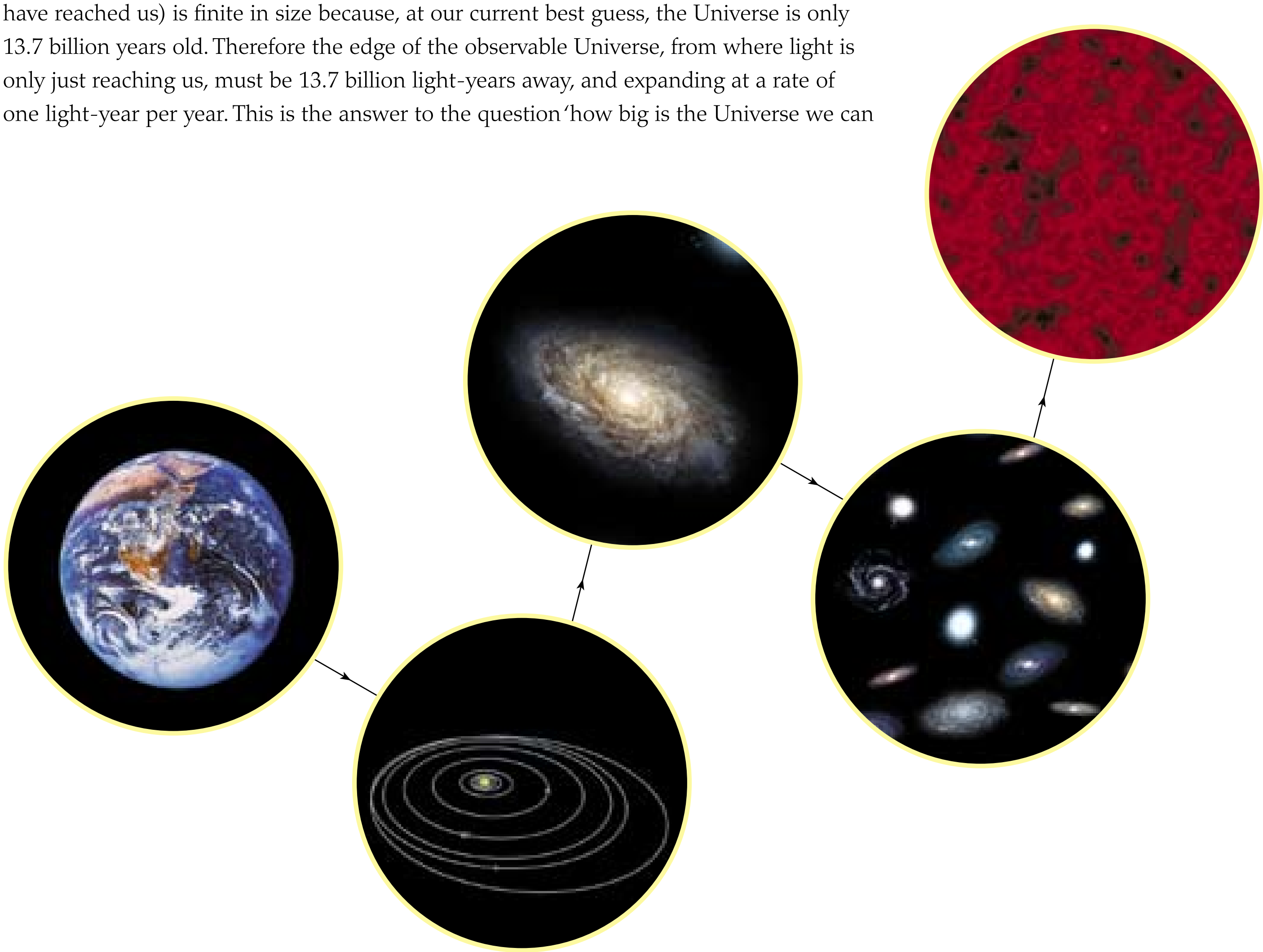


▼ From the invisible to infinity

Fundamental particles; atoms; DNA; cells; humans; planet Earth; the Solar System; our Galaxy; galaxy clusters; and the structure of the Universe. The diagrams show some of the steps on the scale of the Universe.

dimensions of the surface of the football (three if you include time), whereas our Universe is three dimensional (four if you count time).

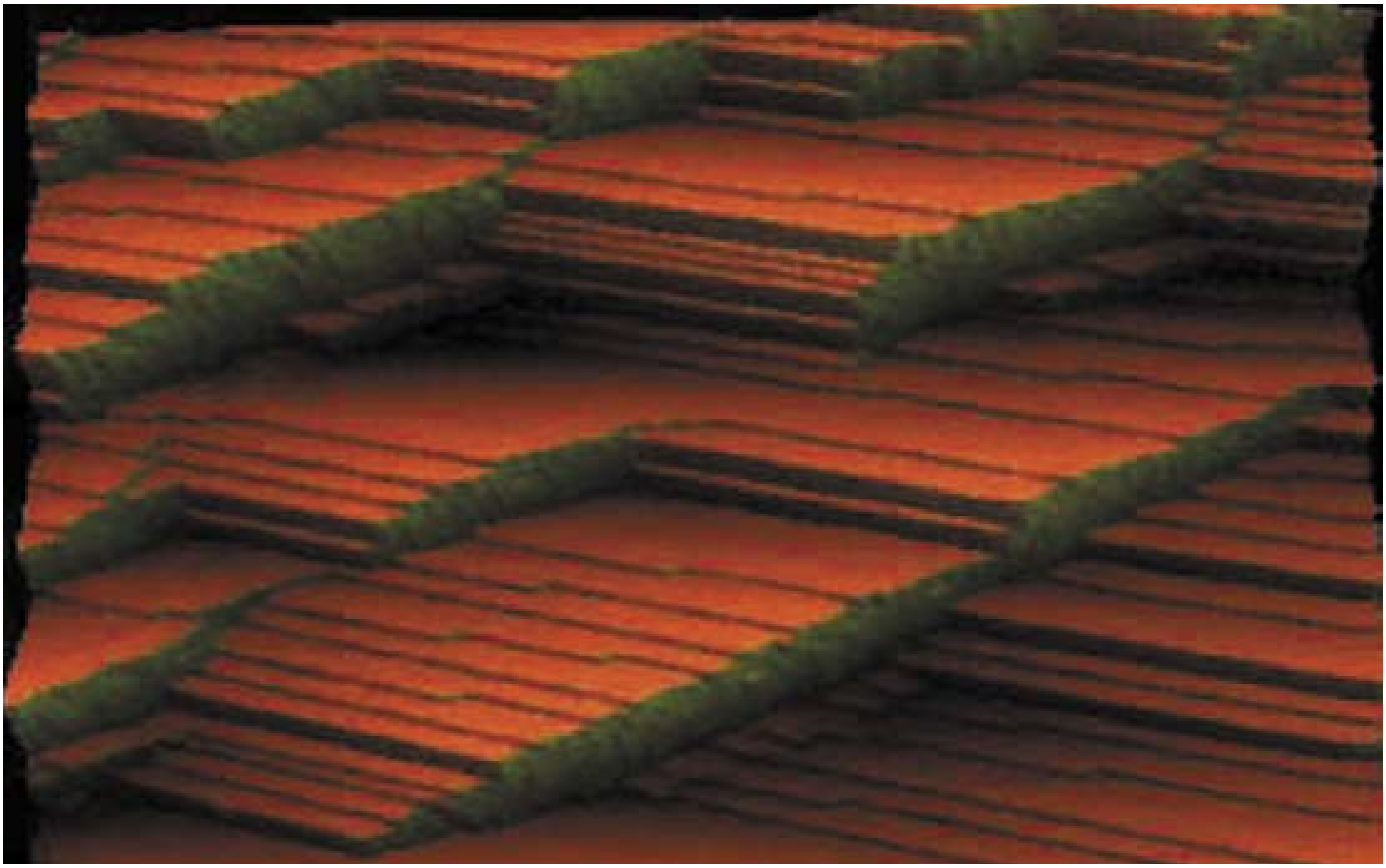
So let us restrict ourselves to questions we can answer scientifically, which means questions we can answer through comparison with observation. We can say for certain that the observable Universe (literally that part of the Universe from which light can potentially have reached us) is finite in size because, at our current best guess, the Universe is only 13.7 billion years old. Therefore the edge of the observable Universe, from where light is only just reaching us, must be 13.7 billion light-years away, and expanding at a rate of one light-year per year. This is the answer to the question 'how big is the Universe we can



see?' As yet, we see no evidence to suggest that we are coming to a full stop; the Universe must be larger than what we can see, and that is all we can say for certain.

The scale of the Universe

Of course, saying that an object is 13.7 billion light-years away is all very well, but can we really comprehend the scale of the Universe? It is possible to appreciate fully the distance between, say, London and New York, or even the distance between the Earth and the Moon (roughly a quarter of a million miles) because this is roughly ten times the circumference of the Earth, and many people have flown for a greater distance than this over their lives – indeed, several airlines give special privileges to those who have travelled more than a million miles in their lifetimes. But how do you really grasp 93 million miles, the distance of the Sun? And when we consider the nearest star, 4.2 light-years (approxi-



▲ **Atomic layers**

Microscope image of atomic layers on the surface of a crystal of iron silicide. The smallest step is only one atom thick.

mately 24 million million miles) away, we are quite unequal to the task ... The galaxies are immensely more distant than this: even the Milky Way's nearest neighbours, such as the Andromeda Galaxy, are over 2 million light-years away.

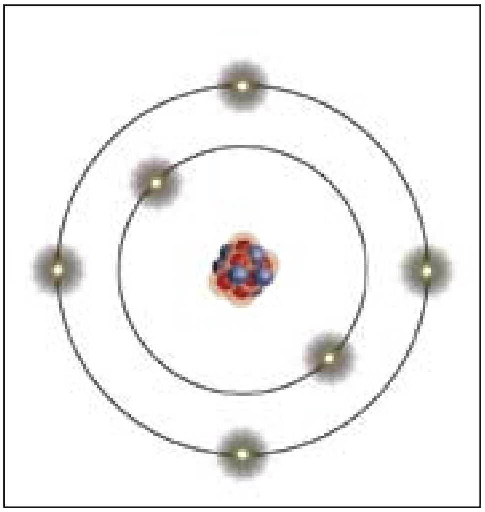
At the other end of the scale, visualizing the size of an atom, which cannot be seen individually with any ordinary microscope, is equally difficult. It has been said that, in scale, a human being is about half-way between an atom and a star. Interestingly, this is also the régime in which physics becomes most complicated; on the atomic scale, we have quantum physics, on the large scale, relativity. It is in between these extremes where our lack of understanding of how to combine these theories really becomes apparent. The Oxford scientist Roger Penrose has written convincingly of his belief that whatever it is that we are missing from our understanding of fundamental physics is also missing from our understanding of consciousness. These ideas are important when one considers what have become known as anthropic points of view, best summarized as the belief that the Universe must be the way it is in order to allow us to be here to observe it.

Meanwhile let us return to question of the size of the Universe. Another way to consider the question is to ask how many atoms there are in the Universe; we find the total number is as large as 10^{79}, or in other words a 1 followed by 79 zeros.

The picture many of us have of each of these 10^{79} atoms is deceptively simple; there are three types of fundamental particle: the proton (carrying unit positive electric charge), the neutron (no charge at all) and the much less massive electron (carrying unit negative charge). Incidentally, it is far from easy to define exactly what electric charge is on the atomic level. It will suffice to think of charge as a property that particles can have, just as they have a size and a mass. Charge always comes in parcels of fixed size which we call unit charge.

An atom then, classically, is a miniature Solar System, with electrons orbiting a central nucleus; the compound nucleus (containing protons and neutrons) carries a positive charge, which is exactly balanced by the combined charge of the orbiting electrons. In our solar system of planets it is the force of gravity which keeps the planets in their orbits around the central Sun, whereas in the atom it is the attraction between the negatively charged electron and positive nucleus that keeps the electrons in their orbits. In passing, we should note that this simple picture can explain much basic chemistry; for example, why it is the outer electrons of atoms that tend to react. They are further away from the nucleus, and therefore they are less tightly held by its attractive force. So the simplest atom, that of hydrogen, has a single proton as its nucleus, and one orbiting electron. The whole atom is therefore electrically neutral: plus one added to minus one equals zero.

This view of the atom, which saw protons and neutrons as solid lumps, prevailed in the early part of the twentieth century, but things are much less clear-cut today. Much of the strange behaviour of extremely small systems can now be explained much better by considering them to be made up of waves rather than particles. This theory is known as wave–particle duality. In addition, experiments have shown that while electrons seem to be truly unbreakable, protons and neutrons are not in fact fundamental – they can be split into even smaller particles known as quarks, which themselves are now believed to be fundamental. No-one has ever seen a quark, but we know they must exist since they have been detected in particle accelerators which have been built to smash protons together at incredibly high speeds. In these experiments protons are seen to fracture, and hence scientists conclude that they cannot be fundamental. Nature abhors a naked quark; quarks appear only in pairs or triplets.



▲ **The classical atom**

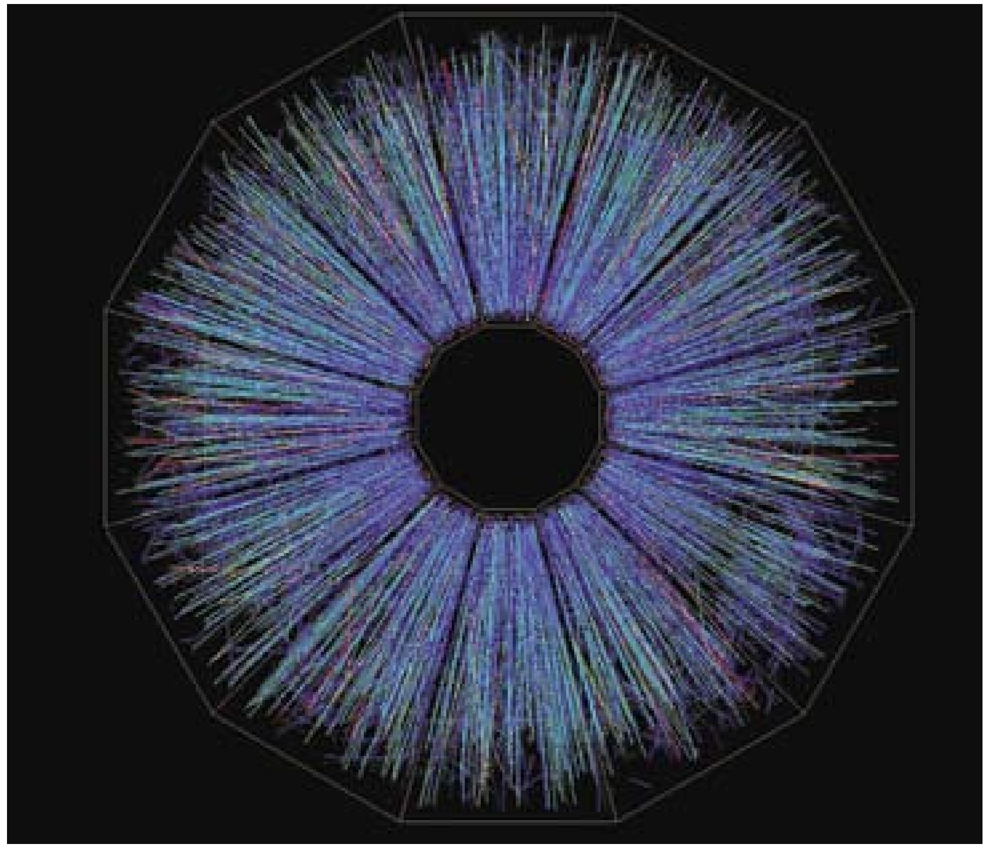
A simple model of an atom comprising a nucleus made up of neutrons and protons, with electrons orbiting like the planets around the Sun, was developed by Neils Bohr in 1913. Despite the advent of quantum mechanics, it is still useful today.

The forces of nature

The reason for this property of quarks lies with an unusual property of the force which normally binds quarks together, known (not without reason) as the 'strong' nuclear force. It is dominant on very small scales, which is why we need such powerful particle accelerators to smash protons apart. Unlike the forces with which we are familiar on larger scales (such as gravity or the attraction between opposite electric charges), the strong force increases with distance. In other words, if we could separate two quarks we would find them being pulled together more and more strongly as the distance between them increases. Eventually, as quarks move apart from each other the energy caught up in straining to pull them apart becomes so great that two extra quarks are produced, the energy being converted into mass. Suddenly we have two pairs of quarks rather than the individual quark we were attempting to isolate. This process means that no experiment ever produces individual quarks, and in the everyday Universe they exist only as 'components' of other particles, such as protons and neutrons which each contain three quarks.

At the huge temperatures in the Universe immediately after the Big Bang, the quarks have enough energy to roam free, and so by understanding the story of the Universe on the largest scales we may come to understand more about the particles which account for the smallest scales. The energy that each particle possessed in the early Universe will

▶ Hunting quarks
At Brookhaven National Laboratory in New York beams of gold nuclei are smashed together at close to the speed of light to recreate the so-called quark-gluon plasma – a state of matter that is believed to have existed ten millionths of a second after the Big Bang. This image of the collision shows about 1000 particle tracks.

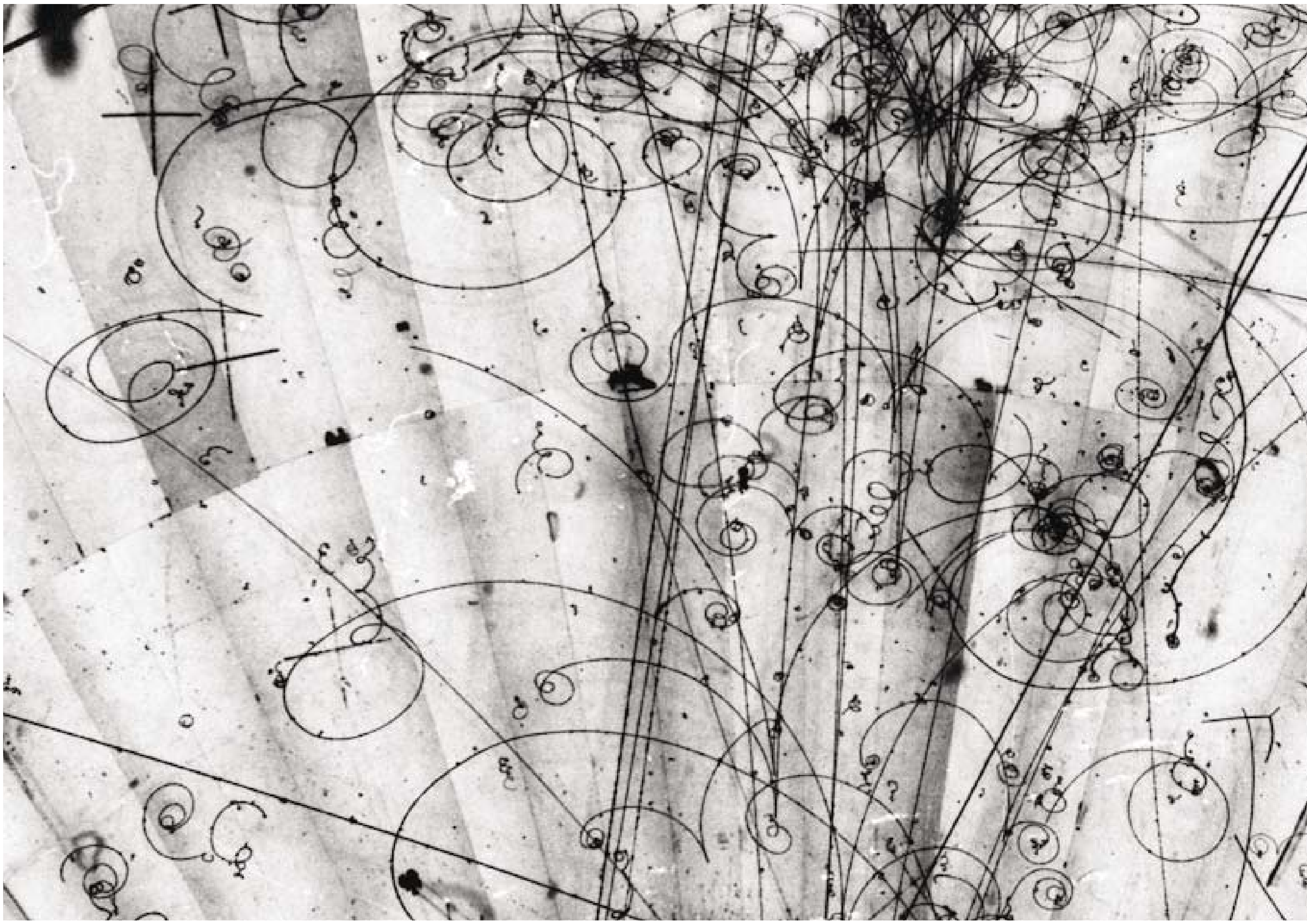


remain far beyond the reach of our particle accelerators; even an accelerator the size of the Solar System would be incapable of producing particles with this enormous energy.

It is a remarkable fact that our present day research into the very small, via particle physics, and into the very largest scales, via cosmology, are so intertwined. To understand the whole Universe, we are dependent on our understanding of the fundamental particles, and our best laboratory for studying them is in the embryonic Universe. A 'hot space' full of these highly energetic fundamental particles is the earliest image we can conjour of our newly-born Universe.

Bigger is cooler

From the first Planck time onwards, this inconceivably small, inconceivably hot Universe began to expand and hence also to cool. The Universe was a sizzling ocean of quarks, each of which had a vast amount of energy, moving at a huge velocity. As a result, there could be no atoms or molecules of the kind we know today, because these are complicated structures, quite unable to survive the disruption of very high temperatures; the quarks were simply too energetic to be captured and confined within protons or neutrons. Instead, they were free to career around in the infant Universe, until they collided with their neighbours. As well as quarks, this early soup of sub-atomic particles also contained anti-quarks – identical twins, but bearing opposite electric charges. (It is now believed that each particle has its equivalent anti-particle, identical in all respects other than its electric charge, which is opposite. The antimatter particle corresponding to an electron is a positron, which bears a positive charge, but is otherwise exactly similar to the electron. The



▲ **Seeing antimatter**

This image shows the creation of pairs of electrons and positrons in a bubble chamber. Charged particles leave a trail of minute bubbles behind them, allowing the eye – or the camera – to track their progress. The photons that provide the energy to create the electrons and positrons cannot be seen, as they have no charge. The electrons and positrons spiral symetrically in opposite directions in this experiment.

concept of antimatter is familiar from science fiction, where it forms the basis of countless highly advanced starship engines, all of which are based on the experimental fact that a collision between a particle and antiparticle results in the annihilation of both and the release of much energy. If a quark were to meet an antiquark in the primeval Universe both would vanish, releasing a flash of radiation. The reverse process also occurred; radiation of sufficiently high energy (certainly at the energies found at this early stage of the evolution of the Universe) could spontaneously produce pairs of particles, each pair composed of a particle and its antiparticle. The Universe at this epoch, then, is composed entirely of radiation which produces pairs of particles, which in turn vanish very rapidly as they collide with each other, returning their energy to the background radiation.

Throughout this whole time, the Universe continued to expand and cool. After the first microsecond (only ten million million million millon million million Planck times), when the temperature had dropped below a critical value of about ten million million degrees, the quarks had slowed down enough to enable them to be captured by their mutual (strong force) attraction. Bunches, each of three quarks, clumped together to form our familiar protons and neutrons (collectively known as baryons), whereas the antiquarks clumped together to form anti-protons and anti-neutrons (anti-baryons).