001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 1

THE BIG QUESTIONS

Physics

Michael Brooks, author of the acclaimed *13 Things That Don't Make Sense*, holds a PhD in quantum physics. He is a journalist and broadcaster, and acts as physics and cosmology consultant to *New Scientist* magazine. He has lectured at Cambridge University, the American Museum of Natural History and New York University, and his writing has appeared in many national newspapers, including the *Guardian*, the *Independent*, the *Observer* and the *Times Higher*.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 2

The Big Questions confronts the fundamental problems of science and philosophy that have perplexed enquiring minds throughout history, and provides and explains the answers of our greatest thinkers. This ambitious series is a unique, accessible and concise distillation of humanity's best ideas.

Series editor **Simon Blackburn** is Professor of Philosophy at the University of Cambridge, Research Professor of Philosophy at the University of North Carolina and one of the most distinguished philosophers of our day.

Titles in *The Big Questions* series include:

PHILOSOPHY
PHYSICS
UNIVERSE
MATHEMATICS

THE BIG QUESTIONS

Physics

Michael Brooks

SERIES EDITOR
Simon Blackburn

Quercus

Contents

Introduction

The beauty of physics is summed up in one simple fact: a child can ask questions that no professor can answer. Indeed, searching out the 'big questions' in physics is rather like looking for hay in a haystack. When it comes to physics, it appears there is no such thing as a small question. A seemingly insignificant query or experiment can often lead to profound insight.

It is a short step, for instance, from asking whether the laws of physics can ever change or be broken to wondering whether there is room for a creator. It doesn't stop there, either. Physics tells us a creator need not be divine; it could be that we live nested within an infinite number of universes, each created by a species only slightly more intelligent than its greatest creation. We may even be destined to become creators of a universe ourselves.

With such big issues at its fingertips, it is small wonder that the most iconic scientists of our generation have been immersed in physics. Albert Einstein became a celebrity almost overnight when his theory of relativity changed our conception of the universe. Carl Sagan's TV programme *Cosmos* remains the most-watched series on public television. Richard Feynman's cool appraisal of the physics behind the *Challenger* shuttle disaster revealed how powerful a working knowledge of the subject can be. Stephen Hawking's work, laid out in his bestseller *A Brief History of Time*, created a thirst for scientific insight in people who had never given the sciences a thought. Only the discoverers of DNA, perhaps, can stand alongside these giants.

And yet, it has to be said, people also tend recoil from physics. If I mention in casual conversation that I am a physicist by training, the announcement is met with a strange mixture of admiration and embarrassment. While expressing awe at anyone who would attempt to understand the universe, many also seem to consider the subject completely beyond them. 'Oh,' they say, 'I never did understand physics.'

If you recognize yourself in that statement, then hopefully this book will change your perspective. Perhaps the best-kept secret in physics is that there is too much there for anyone to understand. This is not a problem, however: this is the root of its allure.

Physics has so much to explore that, once it captures your imagination, it is hard to tear yourself away. The clock on the wall becomes a tease about the elusive nature of time. Sunshine is what results from a beautiful, intricate dance of particles known as nuclear fusion. When raindrops fall to the ground, you can ask yourself a simple 'why?' Exploring the answer will keep you occupied through the longest thunderstorm. The way a sunflower grows speaks of the conservation of energy and how the nature of light has shaped life on earth. Go a step further and ask what light is, and you are peering into something widely considered to be the deepest mystery in nature.

This book is designed to show how simple questions lead to some of the most profound discoveries that humanity has ever made. They encompass the physics you probably didn't learn in class: the real point of the subject; its implications; what we understand about the universe – and what we don't. Carl Sagan once said, 'Somewhere, something incredible is waiting to be known.' Hopefully, that process can begin here.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 8

WHAT IS THE POINT OF PHYSICS?

Impossible questions, unexpected rewards, and the never-ending quest for understanding

The question has bounced around school classrooms for decades. The answer offered usually starts with an apocryphal tale involving the legendary Greek philosopher Archimedes and King Hiero's crown.

Hiero had come to the throne in the Sicilian city of Syracuse. He gave a craftsman a certain quantity of gold to fashion into a crown; when the crown arrived, so did a rumour that the craftsman had substituted some of the gold for silver. Hiero commissioned Archimedes, then in his early twenties, to find the truth.

The story, as related by the Roman writer Marcus Vitruvius Pollio, says that Archimedes realized how to solve the problem when he noticed the volume of water that his body displaced in a bath. Silver, being less dense than gold, would displace less water, and so by dipping the crown in a bath of water it would be possible to tell if there was silver in the crown. In his jubilation, Archimedes rushed down the street naked, shouting 'Eureka': 'I have found it'.

Is this is the point of physics: to answer seemingly unanswerable questions? We are now able to look at our surroundings across an extraordinary breadth of scales. Where we once thought visible matter to be indivisible, we have gone smaller and smaller, down to the atom, and onward to the most fundamental particles, and ultimately to a view where matter is

actually composed of fluctuations in the energy of empty space (see *Are Solids Really Solid?*). The sky was once the limit of our vision; now we know the universe to be so vast that it would take light nearly 28 billion years to cross (see *Am I Unique?*). And, it should not be forgotten, understanding the notion that light has a defined and constant speed is a hard-won triumph of physics too (see *Can We Travel Through Time?*).

We know much of the history of the universe, the nature of matter and the structure of our planet, but perhaps the greatest lesson we have learned is that, whenever we think we have nature figured out, it surprises us again, revealing just how little we actually know. Isaac Newton probably put it best in his memoirs: 'I do not know what I may appear to the world,' he wrote, 'but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.'

'I do not know what I may appear to the world,' he wrote, 'but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.'

ISAAC NEWTON

An alternative to superstition

If there is one aspect of physics' achievements that Newton perhaps appreciated less than most, it was the subject's ability to slice through mysticism and superstition. Newton was a great alchemist and a biblical scholar; he considered his writings on the Old Testament book of Daniel his greatest work. Whenever physics threatened to cast doubts upon spiritual matters, Newton would cringe. 'I have studied these things – you have not,' was his constant retort to astronomers' criticisms of religion. Newton left room for God's work in the mechanism of his 'clockwork heavens' but the march of physics soon displaced the divine hand. When the Emperor Napoleon questioned Pierre-Simon Laplace about his newly published treatise on celestial mechanics, he remarked on the absence of God in the mechanism. 'I have no need of that hypothesis,' Laplace replied. The point of physics, in many ways, is to find what, in the universe, is explicable by a set of laws, and the simpler the laws the better.

Until around 600 BC, civilizations developed technologies but thought little about how to make sense of the world: that was for the prophets and the sages. Then came the Milesians. The city of Miletus, on the west coast of modern Turkey, was home to a mode of thought that would be recognizable to today's scientists as a thirst for real, first-hand understanding. Rather than having the universe's secrets obscured by mystical religious concerns, the Milesians sought laws to explain the phenomena of nature, and came up with theories for the causes of earthquakes, lightning and the structure of the universe, among other things.

The Milesians debated these theories openly, considered how they might be tested, and accepted the results of experiments as the arbiter of truth. Anaximenes of Miletus is credited with performing the world's first scientific experiment. His observations of how the temperature of exhaled breath seems to vary depending on whether the lips are pursed or wide open, led him to conclude that compression causes cooling, and expansion causes heating.

The fact that Anaximenes was exactly wrong here is another lesson in the point of physics. It teaches us that we cannot ever be sure of anything that is 'received wisdom'; accepted theories, and even 'facts' about how things in the universe work, are often proved wrong, and supplanted by new ideas. These, too, are open to falsification. Physics is a process of testing everything – especially those things we most want to be true.

It is for this reason that physics is somewhat devoid of 'scientific saints'. It is not so much a discipline of ideas as a discipline of consensus arrived at through the gathering of experimental evidence. Those who fail to accept the results of experiments – and

WHAT IS THE POINT OF PHYSICS?

PROVIDING THE LAWS BY WHICH SOCIETY RUNS

James Wilson, who played a significant role in the drafting of the American Constitution and became one of George Washington's six original supreme court justices, took the ideas of physics to heart. When defining the role of government in his Lectures on Law, he said, 'Each part acts and is acted upon, supports and is supported, regulates and is regulated by the rest . . . there is a necessity for movement in human affairs; and these powers are forced to move, though still to move in concert.'

Wilson's statement worthy of Isaac Newton – it invokes the same laws of interaction that allowed Newton to deduce how the solar system worked. What's more, the link from Newton to political theory is not a hard one to trace. Newton was inspired by Copernicus, who acknowledged the work of Aristarchus of Samos, who lived in Greece between 310 and 230 BC. Aristarchus was, in turn, inspired by the Greek philosopher, aristocrat and politician Plato. Plato's greatest contribution to civilization is considered to be his Republic, an examination of how best to run a society. But Plato was a distinguished astronomer too – he was the first person to recognize, for instance, that anomalies in the motion of the planets might be resolved by finding some combination of circular motions.

Plato thought physics an excellent training for a politician. Leaders should learn physical sciences such as astronomy, Plato once declared; not because they help in stargazing or navigation, but because they provide an education in the techniques of abstract thought that are essential to leadership. The same skills are still highly valued today: trained physicists are very much in demand outside the laboratory walls – in, finance, business and government.

do not provide good reasons why others should join them on the 'wrong' side of the fence – tend to be given short shrift.

More than the sum of the parts

The physicists Albert Einstein and Richard Feynman provide a suitable illustration of the way physics is bigger than any physicist. Though now venerated as a public icon, Einstein did not die a hero to other physicists. On the contrary, his later life is remembered with a tinge of regret at his ultimate quest. Einstein's best-known work was done early in his career. He made a seminal contribution to quantum theory with the experimental discovery of the photon, the quantum of energy (see *What is Light?*).

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 12

This destroyed the centuries-old view that light must be a wave. Then his special theory of relativity changed our notion of time. His elucidation of the idea that mass and energy are interchangeable (see *Why Does E = mc²?*) was a revelation about the fundamentals of matter. The general theory of relativity rewrote Newton's gravitational work after nearly four centuries of acceptance (see *Why Does an Apple Fall?*).

After that, though, Einstein's views grew irrelevant to physics. The quantum revolution changed the face of the subject, but Einstein refused to accept quantum theory as a useful way to describe the universe. He spent his later years working, to no avail, on a theory that would unite electromagnetism and relativity and render quantum theory an unnecessary innovation. The number of physicists who would work with him and support him dwindled throughout his life.

Richard Feynman is perhaps the second most famous physicist after Einstein. He is a great popularizer of the subject, a great and innovative thinker, and – most significantly of all – remains a great hero to those working in the field. Feynman never reached Einstein's dizzy heights of achievement, but he did more than most, contributing to the creation of quantum electrodynamics, or QED, a theory that describes the interactions of light and matter (see *What is Light?*). It is widely feted as our most successful theory of physics.

> *'The first principle is that you must not fool yourself – and you are the easiest person to fool.'*
> RICHARD FEYNMANN

One of Feynman's greatest strengths as a physicist was his ability to listen to the convictions of his peers, bow to the law of evidence, and admit that he was always working from a position of ignorance. He famously said that, 'The first principle is that you must not fool yourself – and you are the easiest person to fool.' His unwillingness to fool himself is summed up in his appraisal of the theory that became Einstein's downfall. 'I think I can safely say that nobody understands quantum mechanics,' he wrote in *The Character of Physical Law*. 'Do not keep saying to yourself, if you can

possibly avoid it, 'But how can it be like that?' because you will get . . . into a blind alley from which nobody has yet escaped. Nobody knows how it can be like that.'

This is the reason the older Einstein is not revered by physicists, and Feynman is. While Einstein led himself into a blind alley, Feynman admitted his limited understanding, and followed others as they made forays into new territory. This is another component of the point of physics: progress by building on the achievements of others. As Newton put it, 'If I have seen further it is only by standing on the shoulders of giants.'

> *'If I have seen further it is only by standing on the shoulders of giants.'*
> ISAAC NEWTON

Thanks to quantum theory, physics has even taken the extraordinary step of defining some limits for itself. The Heisenberg uncertainty principle (see *Is Everything Ultimately Random?*) sets in stone the fact that there are limits to what physics can tell us about a system.

A humble discipline

When we examine the equations that govern the motion of an electron, say, we can see how they tell us its momentum, or its velocity. There is no means by which they can tell us, precisely, about both the momentum and the velocity, however. The two can be found to only a finite precision.

Werner Heisenberg saw the practical side of this: there are limits in what our experiments can reveal. Bounce a photon of light off the electron, and you can infer its position, but the photon will have imparted some momentum to the electron, too. Thus the act of determining the position of the electron creates an uncertainty in the value of its momentum. Conversely, a measurement of momentum will always create an uncertainty in a particle's position. Whether you look at theory or experiment, there are strict limitations to what we can find out. Physics, in many ways, is a humble discipline. But there's plenty to be humble about, as the physicists behind the atomic bomb will testify.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 14

If you had posed the question 'what is the point of physics?' to Western governments after the Second World War, you would have been greeted by disbelief that you even had to ask. Physics was everything, as the war had shown. Physics had given us fantastic technological innovations: radar, computers, the atomic bomb, and, of course, televisions and microwave ovens. Physics was set to be the driver of economies, and the protector of nations. Pose the same question to physicists, however, and you might have got a rather more subdued response.

Immediately after the first test of the atomic bomb in New Mexico, the Harvard physicist Kenneth Bainbridge turned to Robert Oppenheimer, the project leader. 'Now we're all sons of bitches,' he said. Oppenheimer was dealing with his own mixed emotions: decades later, he admitted they all knew at that moment that the world would never be the same. And yet, Oppenheimer said, put in the same situation, he would do it all again. 'If you are a scientist, you cannot stop such a thing,' he said in his retirement speech in 1945. 'If you are a scientist, you believe that it is good to find out how the world works . . . that it is good to turn over to mankind at large the greatest possible power to control the world.'

The world in our pocket

Is this the point of physics: to gain control over the world? It is true that physics – or at least the industrial application of physics – has created the modern world. If our age can be defined by one thing, it is probably the microelectronics revolution: television, computing, the Internet, and mobile communications, to mention but a few aspects. All of it was built on the back of physics. To be more specific, it was built on the back of silicon technology. During the Second World War, the developers of radar worked to create ever-purer crystals of silicon and germanium for the equipment. Physicists – above all the ones employed by Bell Labs in the USA – continued that development after the war, learning how to turn them into 'semiconductors' and incorporate them into technologies that had previously required inefficient and bulky valve amplifiers. By 1952, the first silicon-based electronics products had hit the market: low-power and highly portable

devices, such as hearing aids and pocket radios. A year later, the first transistor-driven computer appeared. Shortly after that, people started to refer to the concentration of electronics companies in a small area of northern California as 'Silicon Valley'.

It is not hard to see the impact of physics on our lives. Lasers provide a specific example. Lasers also came from Bell Labs, and stemmed from wartime research into radar technology. Since their invention in 1957, they have become ubiquitous in everyday life. CD and DVD players, fibre-optic communications systems such as the telephone network, supermarket checkout scanners, eye surgery and laser printers are just a few of the applications.

So, is the development of technology the point of physics? Not at all. The technological revolutions of the 20th century came about as a result, ultimately, of the discovery – or invention, if you prefer – of quantum theory. That was the result of trying to unravel things no one understood, such as why the spectrum of radiation emitted by an oven at 100 celsius was the same as the spectrum of radiation emitted by anything else at 100 celsius, rather than specifically trying to invent new devices.

In essence, our modern electronic technologies, come from quantum theory, which came from thermodynamics, the study of heat. That arose from the study of gases – and so on. Physics is a self-sustaining chain reaction: every discovery provokes another set of questions, which provoke new discoveries. As George Bernard Shaw once said, science never solves a problem without creating ten more.

A never-ending story

There is no end of questions in sight. Physicists used to be fond of saying their work was done. In 1894, the American physicist Albert Michelson announced that, 'The most important fundamental laws and facts of physical science have all been discovered, and these are now so firmly established that the possibility of their ever being supplemented by new discoveries is exceedingly remote.' Within a decade, we had the twin revolutions of relativity and quantum theory.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 16

In 1888, the astronomer Simon Newcomb had announced the end of astronomy: there was little left in the heavens to discover, he suggested. Newcomb was wrong too. Our view of the cosmos has probably changed more radically since Newcomb's time than it did in the thousands of years of scientific discovery that took place before he was born. Although the major breakthroughs of the last century showed us where we came from, outlining the entire history of the universe, the hubris is gone from our world view; with the discovery that most of the universe is in a form unknown to science, physicists now appreciate that they have got to grips with only a tiny percentage of the universe.

There is, it has to be said, one end in sight: the theory of everything. If physics began with the Milesian quest for the laws governing natural phenomena, it will (theoretically) end with the discovery of just one law: the ultimate description of the universe. This 'theory of everything' will reduce all the particles, the forces that govern their interactions and the space and time in which their existence plays out, to a single unified description (see *Is String Theory Really About Strings?*).

At the moment, we are far from achieving that goal, but here, perhaps, we have found the true point and the essence of physics: to discover the span of our ignorance, and to do what we can to reduce it. Sometimes, as with the atomic bomb, there is a price to be paid for this journey of discovery. Sometimes, as with the development of quantum mechanics, we reap great practical rewards from it. But most of the time, physicists will tell you, physics is simply about the thrill of discovery – and then discovering that our discoveries have made the world more interesting, not less. As the poet John Dryden said, 'Joy in looking and comprehending is nature's most beautiful gift.'

WHAT IS TIME?

Progress, disorder and Einstein's elastic clocks

Deep in your brain there lies a lump of tissue called the striatum. This assortment of neurons is, to the best of our knowledge, the only dwelling place of time. It accumulates the first record of the moments of your life, and provokes your sense that your childhood was a tumbling assortment of significant and fascinating moments, while adult life hurtles by too fast to be properly appreciated.

You shouldn't set too much store by these sentiments, though. The striatum's gift is actually to create an impression – perhaps even an illusion – of time passing. The problem is, its measure of time depends on what is going on in your conscious mind. Every time you perform a conscious task such as putting the kettle on, the various electrical circuits in your brain spike in unison. The striatum records this simultaneous signalling and starts to note the subsequent patterns of electrical signalling from areas such as the frontal cortex. Your notion of how much time has passed before the kettle boils is nothing more than a measure of the accumulated electrical signals.

That's not so bad at home, where you can calibrate it with a glance at the kitchen clock. But as soon as you are denied access to clocks, things go awry. When, in the early 1960s, the French geologist Michel Siffre took off his watch and lowered himself into a dark cave for 60 days, his perception of passing time unravelled.

folio position OK (master page item)

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 18

By the end of the experiment, what Siffre thought was an hour was often four or five. Drugs such as valium, caffeine or LSD will send your sense of time similarly awry. As will your memory.

We often think busy times make life flash by, but experiments show that's only true *while* you're busy. Afterwards, when you reflect on your existence, your busy periods will seem much longer. That's why your childhood now seems to have been a series of long, golden summers – life was exciting when you still had so much to experience, and your brain thinks that those heightened signalling levels must correspond to huge stretches of time. Your grip on the passage of time, then, is as precarious as you may always have suspected. But it turns out that our problems with the perception of time are as nothing compared with our problems with the notion of time itself.

Universal time

You might think we ought to have a handle on time by now. After all, time is a universally understood concept – every human culture knows about it, talks about it, feels it. And we have been thinking about what it means for millennia. In 350 BC, Aristotle, for instance, wrote a work called *Physics*, which included one of the first attempts to grapple with the notion of time.

> *'First, does it belong to the class of things that exist or to that of things that do not exist?'*
>
> ARISTOTLE

Aristotle's work on time begins with a question. 'First,' it says, 'does it belong to the class of things that exist or to that of things that do not exist?' Here in the second millennium AD, that is still an open question. If our minds are fooled by the passage of time, that may be because time itself is an illusion. From the Greeks to modern- day physics, the main conclusion about time has remained constant: it is, at the very least, about change. Through time, one thing changes into another.

But while Aristotle's Greek peers were obsessed with the circle as the most fundamental concept in the universe, meaning that time must flow in cycles, modern physics is focused on linear

processes: beginning to end, big bang to cosmic shutdown. With time, that translates into an overwhelming sense of time's arrow: in our modern view of the universe, time moves irreversibly forward. Eggs break, and cannot be unbroken. Clocks wind down, and do not spontaneously wind up.

This process of change, in which systems move irreversibly into disorder, is known as the thermodynamic arrow of time. It arises from one of the most fundamental laws of physics: the second law of thermodynamics. This states that, as a whole, the universe is caught in a process of unravelling order. Entropy, a measure of the disorder in a system, is always increasing.

Order and disorder

The arrow of time might arise from a variety of sources. The 'cosmological arrow of time', for example, cites the creation of the universe as a move away from a special, high entropy state where everything was neatly ordered. It is rather like handing a fully solved Rubik's Cube to a curious child; as time progresses, the universe moves to an ever-more disordered state, just as the neat order on the faces of the Rubik's Cube will give way to a messy jumble of colours. While some things, such as galaxies, appear ordered, with structures that are often intricately beautiful, the order of the universe as a whole is decreasing. The end will come when there is no more disorder to be created; or, as Lord Kelvin put it, when the universe has reached 'a state of universal rest and death'.

Our familiar arrow of time could equally result from quantum theory. In one (probably the most popular) school of thought, quantum systems undergo an irreversible 'collapse' when they are measured. This originates from the remarkable ability of a quantum object such as an atom to exist in two entirely different states at once. It might, for example, be spinning clockwise and anticlockwise at the same time. When the measurement is made, however, that double state is forced to become one or the other: the measured atom will be found to be spinning clockwise or anticlockwise, and will not spontaneously revert to the state of doing both.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 20

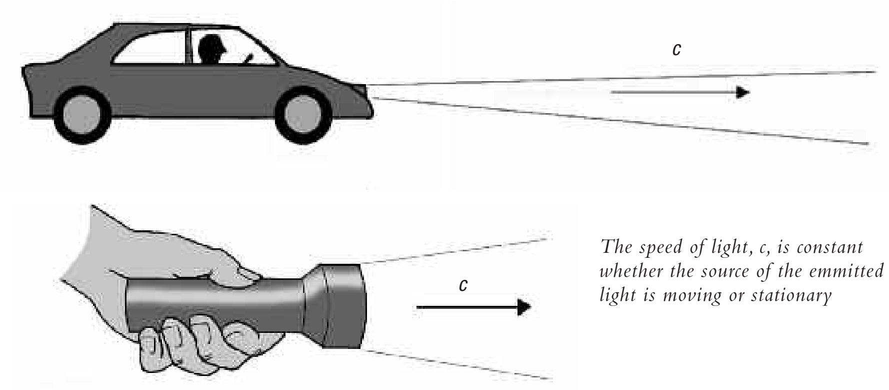
There is a problem with these descriptions of time's arrow, however. They get us nowhere because they require the concept of change. And change, as Aristotle noted, is a marker of time passing. Through considerations of the arrow of time we are really no further forward in defining time. All we have is a putative explanation for the direction it appears to take. And even that has been undermined. Time's arrow might be part of our individual experience, but we have no reason to believe that makes it real. Worse still, we have good reason to believe it isn't.

A stretch in time

We have Albert Einstein to thank for this alarming insight: it lies at the heart of his special theory of relativity. Einstein was relatively unknown when he published his ideas in 1905. Special relativity was a revolutionary work, dismissing in a single stroke the popular and long-lived concept of the aether, a kind of ghostly fluid that fills all of space and provides a background through which electromagnetic fields such as light could move.

It is worth mentioning at this point that while, as the late Carl Sagan once said, extraordinary theories require extraordinary evidence, special relativity is one of the few such theories where extraordinary evidence has been found to back it up. What you are about to read may seem absurd, but there is every reason to take it seriously.

The central point of special relativity is that the laws of physics work the same for everyone, regardless of how they are moving through the universe. The most important consequence of this is that the speed of light is a constant, universally known as c. If you were to measure the speed of the light emitted from the headlights of a vehicle travelling towards you at 100 kilometres per hour, the speed of the light would be c, not c plus 100 kilometres per hour (62 mph). The speed of light does not change depending on the relative motion of the emitter and observer. The extraordinary upshot of the constancy of c is that, when conditions require it, everything else does change – and that



The speed of light, c, is constant whether the source of the emmitted light is moving or stationary

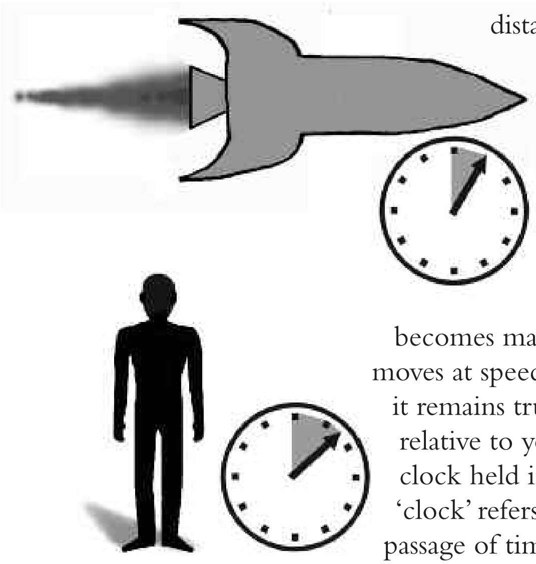
THE CONSTANT SPEED OF LIGHT

includes time. The passage of time is as flexible an affair in the real, physical world as it is inside your mind.

Let's imagine a scene where you are standing 100 metres from an intersection controlled by traffic lights. You are equipped with a stunningly accurate stopwatch, a metre rule and lightning reflexes. The light changes to red, and you are able to measure the time it takes for the first pulse of red light to travel the length of your metre rule. At that moment, a car passes you, travelling towards the intersection at one 100 kilometres per hour. The passenger in the front seat has the same skills and equipment as you, and makes the same measurement: the time taken for the light to travel the length of the ruler.

You have both measured the speed of light, and Einstein insists that you must both get the same result. But as the car moved past you towards the traffic light, the metre rule within it also moved past you. By the time the light reached the end of the ruler in the car, the far end of the ruler was closer to the traffic lights, and so the light had to travel less distance compared with yours. The passenger in the car should measure light as faster, completing a metre in less time. How then, can you both get the same result? The answer has to do with the passage of time in different situations. Compared with your clock, the clock in the moving car runs slow. So, although the light apparently had less

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 22



TIME DILATION

distance to travel, the time measurement was larger than yours, cancelling out the effect.

This is not a sleight of hand where a combination of illusions leads to you getting the right result. The effect, known as time dilation, only becomes markedly noticeable when the clock moves at speeds close to the speed of light, but it remains true that a clock that is moving relative to you really will run slower than a clock held in your hand. And the word 'clock' refers to anything that can mark the passage of time. Dissect that statement, and you'll find that all kinds of disturbing implications emerge.

Ageing relatives

Let's start with something that is just about conceivable. Take a lump of polonium, a radioactive material discovered by Pierre and Marie Curie around 200 years ago. One form of polonium, polonium-209, has a half-life of about 100 years; that is, after a century, half of its atoms will have emitted a burst of radiation and transmuted into more stable atoms.

If the Curies had taken two identical lumps of this material when they discovered it, and left one in their Paris lab while shooting the other one on a round trip into space at 0.99 of the speed of light, returning to Earth today, we would notice something remarkable about the amount of radiation they were giving off. The lump that stayed in Paris would lose half of its radioactive polonium atoms in the first 100 years. During the second century, it would lose half of the remaining half. So, after 200 years, 75 per cent of the polonium would have decayed. The thing is, its twin, the lump that had rocketed into space and back would only have lost 20 per cent.

That is because the motion relative to earth at 0.99 the speed of light (setting aside practical issues such as acceleration, deceleration and turning round) slows time for this lump. Its 'clock', as measured by the rate at which its atoms experience radioactive decay, is running at only 14 per cent of the speed of its twin that never left the planet. That is why so many of its radioactive atoms remain intact. This, perhaps, is hard enough to swallow. But now for something truly inconceivable.

Let's allow Pierre and Marie Curie to guard the two lumps of polonium. Pierre will accompany one lump on that same return trip into space, while Marie remains in Paris with her lump. The scientists' bodies have internal clocks, too: as with the polonium, their atoms change with the passage of time, creating a heartbeat, for instance, and cells that shut down after performing a certain number of divisions – a phenomenon that biologists believe to be the root of ageing and death.

Turning a blind eye to the likely catastrophic effects of the radiation, the atoms – and thus the cells and the heartbeat – in Pierre's body will run slow compared to Marie's, just as the polonium's radioactive decay runs slower than on earth. When Pierre returns, 200 earth years later, Marie is long dead, but Pierre's body has aged only 28 years. One immediately obvious conclusion from this is that, given the right resources, time travel into the future is entirely possible. But it is a short step from this point to the astonishing revelation that Einstein's special theory of relativity does away with the notion of some common future anyway. And neither is there a common present or past.

In search of lost time

You might claim, as you stand looking at the traffic lights, that you saw two events happen simultaneously. But as we have seen, the passenger in the car has a clock that runs at a different speed. The information they gain about the timing of those two events could well be different. Worse, you might see two events, A and B, happening at distinct times, with B following A. Depending on how your relative friend is moving, however, they could see A

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 24

follow B. That is potentially catastrophic: if you think A caused B, how is that explicable to someone who saw B happen first?

Past, present, future, simultaneity, cause and effect – nothing is universal. When it comes to time and the processes it governs, you and your striatum really are on your own. There is a simple answer to all this confusion, however, and it is an answer that is appealing to many physicists and philosophers. We could do away with the very notion that time exists.

It is an argument that harks back to the 17th century. Newton, whose Christian faith required that space and time reflect the character of God, considered time to be a real entity, an absolute that moves on independently of everything in the universe. But his great rival Gottfried Leibniz believed time to be a human construct. All we can do, Leibniz said, is describe how the positions of things in space relate to each other, and how that relation evolves. It is useful that a clock's pendulum swings back and forth and the clock's hands circulate around the dial in response, for example, but that doesn't mean the clock is measuring something that actually exists. Time, in this view, comes out of our desire to make sense of the world, but it is no more than a useful means of orientation. It is a shorthand, like the spatial concept of 'up'. 'Up' means a certain direction when I am stood in London, but the same direction is actually 'down' in Sydney.

This link is slightly more than a convenient illustration. When Einstein published his general theory of relativity (the 'special' in 'special theory' refers to a special, i.e. particular case, not a special significance), he postulated a bond between time and space. Time, he said, is just one of four dimensions to the universe. The other three are the familiar ones in which you move your physical body: up and down, across, forwards and backwards. The only difference is that, while we conscious creatures can choose how we move through the spatial dimensions, we have no control over our movement through time.

Stretching space and time

Einstein's four dimensions of space and time – together known as space–time – can be thought of rather like a piece of fabric that can be distorted, bent, folded, twisted and even torn by anything within them that has mass or energy. From this foundation, general relativity has equipped us with equations that describe the features of the cosmos with unprecedented accuracy, allowing us to find out how the universe works, send spacecraft to distant destinations and create the array of global positioning satellites that tell us where on earth we are. But perhaps most intriguingly of all, the pliable nature of Einstein's four-dimensional fabric hints at the origin of time.

Your mass distorts space–time very little. The mass of the sun distorts it much more – according to general relativity, this distortion is the root of the gravitational attraction that keeps our planet in orbit. Even more powerful is the distortion that is brought about by a collapsed giant star: a black hole. And it is here that we glimpse the true power of Einstein's work.

The enormously strong gravitational field of a black hole means that there is a spherical region close to its centre where the velocity required to move away from the black hole is greater than the speed of light – an impossible velocity to achieve. Nothing, including light, can get out of this region, and so we cannot gain any information about anything that goes on beyond its boundary. Hence its name: the event horizon.

At the event horizon, time dilation is infinite. Somebody watching from a safe distance as you fall towards the event horizon would see your movements slow down then freeze as time runs infinitely slowly for you compared to the observer. Only in the observer's infinite future would you reach the event horizon, so you never actually disappear from view. Your experience, on the other hand, would be hugely dramatic. Your body is extremely unlikely to survive the enormous gravitational forces, but if you did survive you would eventually encounter what, according to relativity, is a breakdown in the very fabric of

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 26

space–time. This 'singularity' at the centre of a black hole occurs as the distortion becomes infinite. Here, we reach the limit of the known laws of physics – beyond this point, they no longer apply.

The moment when moments began

Though it is commonly associated with destruction, the singularity is also thought to be the key to creation. In the early 1970s, Roger Penrose and Stephen Hawking adapted the mathematical notion of the black hole singularity to explain the origin of the universe. In a black hole everything disappears into the singularity. Reverse the mathematics of the process, though, and the singularity could give birth to the very fabric of space–time. For more than three decades this has been seen as our best description of the big bang, the origin of time itself.

If general relativity sheds some light on where time comes from, it still does not tell us a great deal about what time is. What's more, impressive as Einstein's formulations of the character of space and time are, we know that special and general relativity are not the final answer.

If the singularity shows us anything, it is that, while general relativity works remarkably well in many scenarios, it offers no satisfactory explanation for the most extreme phenomena of our universe. A more complete description of the cosmos and how all its contents (including the centres of black holes) behave – a theory often referred to as 'quantum gravity' – still eludes us. And, as it turns out, the nature of time is right at the heart of the problem.

Quantum gravity has to work relativity's notions of time into quantum theory, our best description of how the microworld of molecules, atoms and subatomic particles behaves. But quantum theory takes little note of time. In the standard formulation of the theory, you can't ask questions about how long a process takes, for example. Then there's the problem that quantum theory tells us that most of the subatomic particles exist independently of the direction of time. Just as they can spin clockwise and anticlockwise at the

same time, their quantum states can evolve forward and backward in time. Researchers are even learning to do quantum experiments where information seems to come from the particles' futures. What's more, special relativity tells us that massless particles, such as photons and the gluons that bind nuclei together, travel at the speed of light and do not even experience the passage of time.

The great physicist John Wheeler once said, 'Time is nature's way to keep everything from happening at once.' He would have said it with a twinkle in his eye, knowing full well that the apparent simplicity of time belies its true nature. Saint Augustine was more honest when he said, 'What then is time? If no one asks me, I know what it is. If I wish to explain it to him who asks, I do not know.'

Despite all our scientific achievements since Augustine, time remains an enigma, possibly the biggest question facing physicists today. But if time is an illusion, it is at least a useful one. Our interpretation of its consequences – our memories of the past, our existence in the present and our hopes for the future – lie at the core of the human experience. Or that's what your striatum wants you to believe.

WHAT HAPPENED TO SCHRÖDINGER'S CAT?

Quantum physics and the nature of reality

It was 1925, the heyday of Buster Keaton and Charlie Chaplin. The world was getting excited about The Gold Rush*, hailed as Chaplin's finest film to date, coming out next month. And poor Wolfgang Pauli, a physics student based in Hamburg, Germany, was depressed. 'Physics at the moment is again very muddled; in any case, for me it is too complicated,' he wrote to a colleague. 'I wish I were a film comedian or something of that sort and had never heard about physics.'*

Pauli was right: physics was muddled. No one understood what the newly formed quantum theory was all about. Experiments dictated that energy must be split into indivisible packets or quanta, but no one could say why. Then, just a few months later, the Austrian physicist Erwin Schrödinger cleared the confusion. It happened during a trip into the Swiss mountains with a woman who was not his wife, and ended with his questioning the fate of an imaginary cat. The creature quickly became the most famous animal in science. The story of Schrödinger's cat has the weirdness of quantum theory running right through it, and its appropriately enigmatic nature remains intact to this day.

The source of Schrödinger's breakthrough lay in the work of a French physicist called Louis de Broglie. In 1923, de Broglie put together relativity, generally the physics of the very largest

scales of distance and speed, and the nascent quantum theory, the physics of the very small. The outcome was a simple equation. Every moving particle, de Broglie said, could equally well be described as a wave. Every wave could be described as a moving particle. Einstein, when presented with the work, pronounced it 'quite interesting'. Two years later, however, Schrödinger showed it was much, much more than that.

Erwin Schrödinger worked out the mathematical implications of de Broglie's formula during a Christmas holiday in 1925. Leaving his wife in Zürich, Schrödinger took his mistress off to a chalet in the Swiss mountains. It was not unusual behaviour for him – he and his wife seemed to come to several 'arrangements' through their marriage. Whatever went on, the trip was obviously inspiring. Schrödinger came back from the mountains with what is now known as the Schrödinger wave equation. This describes how a quantum particle behaves when it is considered as a wave.

The Schrödinger equation provides a way of understanding where quantum states come from. Take the Bohr model of the atom, for example, where an electron circling the nucleus can only have particular energy states. Schrödinger's equation gives a way of working out what those 'quantized' energies are: the electron is stable only when its wave completes a whole number of oscillations during its orbit.

This was a revelation to physicists, who had no proper justification of the quantized energies. But the equation also gives a way of working out how the energy, say, of an electron will evolve over time in a particular situation. It can equally well give us the particle's position, or its momentum, or how the quantum states of two interacting particles will end up. It was hailed as a masterstroke. There was only one problem.

No one could agree on what the wave equation actually meant. Did it mean the particles were really waves? Schrödinger believed – or rather hoped – so. Einstein stood with him. But

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 30

others disagreed. The University of Göttingen physicist Max Born, for instance, showed that solutions of the wave equation might give nothing more than probabilities. The probability of finding a particle in a particular space, say, or the probability that a particle will have a certain momentum.

In this view the equation was a guide to what we might find out about the quantum system under inspection, but had nothing to say about what the nature of the system actually was. In other words, it did not give us a description of the quantum object, only a description of what we could know about it. Philosophically, this was a nightmare. Einstein hated it, as did Schrödinger.

Positive thinking

Niels Bohr, on the other hand, loved it. Bohr was based in Copenhagen, where he ran an institute sponsored by the Carlsberg brewery. He was a 'positivist': his philosophy said that it was meaningless to talk about something's objective properties because you could only ever access knowledge about it through subjective measurements. Those measurements will always impose restrictions on what we can know.

The ultimate reality behind Schrödinger's wave equation was neither wave nor particle, Bohr felt, and so could not be described in any terms we can deal with. His answer was to assume that nothing exists until it is measured. But once a measurement is made, the type of measurement will determine what we see. If you use an instrument that detects something's position in space, for instance, you'll see something that has a definite position in space – the entity that we call a particle.

Einstein would have none of this 'Copenhagen interpretation' of quantum theory. His great work, relativity, had been built specifically to create a theory that was independent of the observer. The central theme of relativity was that the laws of physics should be the same, whoever is working them out. The notion that the physical nature of the universe was dependent on how we looked at it offended his sensibilities deeply.

Einstein's problem lay in the fact that describing quantum objects using a wave equation meant that, like waves, they could interfere with one another. When two waves interact, they produce a 'superposition', which is the sum of the waves at any point. Where two crests coincide, the superposition is larger than both. When two troughs coincide, the wave trough deepens. If a crest and a trough coincide, the result is flat.

How does this apply to quantum particles? Schrödinger's wave equation says that, in the right circumstances, they exist in a superposition of different states. Thus an electron circulating in a ring of metal can be circulating clockwise and anticlockwise at the same time. A photon of light can be polarized – that is, have its electric field oriented – in any number of directions at the same time. A radioactive atom, which decays via a quantum process, can be in a superposition state of 'decayed' and 'not decayed'. Though it seems nonsensical, this is what the theory states.

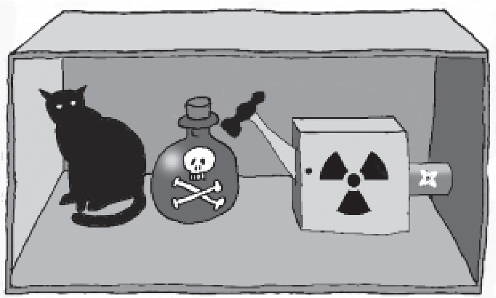
Which is why Einstein and Schrödinger said there must be something missing from the theory. And, to hammer his point home, Schrödinger came up with the cat. 'One can even set up quite ridiculous cases,' Schrödinger wrote in a 1935 journal article. 'A cat is penned up in a steel chamber . . .' Schrödinger went on to describe this 'ridiculous' case in some detail, unwittingly creating the touchstone for future interpretations of quantum theory.

Cat in a box

In the closed steel chamber with Schrödinger's cat is a tiny piece of radioactive material and a Geiger counter. At any moment, there is some probability that the radioactive material will emit a particle, thus triggering an electrical current in the Geiger counter. But Schrödinger had the Geiger counter rigged up to release a hammer that, on sensing a radioactive emission, would smash a flask of hydrocyanic acid, releasing vapours that would kill the cat.

According to Schrödinger the quantum description of the entire system, including all the atoms that make up the cat, 'would express this by having in it the living and dead cat (pardon the

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 32



SCHRÖDINGER'S CAT THOUGHT EXPERIMENT

expression) mixed or smeared out in equal parts.' The logic is sound. The indeterminate nature of the radioactive atom, in a superposition of 'decayed' and 'not decayed' can also put the cat in a superposition of dead and alive.

The kicker comes when the issue of measurement is brought to bear. Bohr had said that there is no definite reality until a measurement is made, because the choice of measuring instrument determines which facet of the system – wave or particle, for instance – the observer will see. So, in Bohr's view, the act of opening the box and observing the state of the cat would force it to be alive or dead.

This was what Schrödinger found so ridiculous: how can the act of observation change such a fundamental property of a cat? It must be one thing or the other; Bohr was being fooled in the same way that a blurred photograph can give an impression of fog, he said. 'There is a difference between a shaky or out-of-focus photograph and a snapshot of clouds and fog banks.'

By this time, though, the interpretation of quantum theory was already a matter of public debate: Einstein and Bohr had a famous exchange in 1927, at the fifth Solvay Conference in Brussels. Einstein challenged Bohr with a series of thought experiments. Imagine such and such a situation, he would say: how can the observation, or the interaction with the apparatus cause a superposition to resolve into one state or the other?

Waves and bullets

The eventual outcome culmination of this arguement was a new version of an old experiment: the famous 'double slit' experiment. In 1801, Thomas Young overturned Newton's particle view of light by shining light at a screen scored with two slits. Young

observed an 'interference' pattern, which can only be explained through superposition of waves. The quantum version asks what happens when you reduce the light intensity so far that quantum theory kicks in. When there is only one bullet, or 'photon' of light in the experiment at any one time, there can be no interference, surely?

In Bohr's view there could – as long as no one was looking to see which slit the photon travelled through. To Bohr, the light is neither a wave nor a particle – those are names that we give something whose properties we have measured. According to Schrödinger's wave equation, the photons of light go through both slits. Despite being a single particle, each photon is 'smeared out' as a wave, effectively having two independent existences as it passes through the slits. As long as no one measures the path the light takes, it takes all available paths.

You might think that this is all wordplay – abstract thought experiments whose weirdness will disappear once the experiments are carried out in the real world. You would, to Bohr's delight, be wrong. We didn't find that out for sure until relatively recently. The first double slit experiment with only one particle in the apparatus at any one time was only carried out in the 1970s. But it worked: despite being faced with two slits, a succession of electrons gradually built up an interference pattern on the screen beyond the slits.

And spookily, when an instrument was placed in the experiment to measure which slit the electron went through, the interference pattern disappeared. In other words, measurement made it manifest as a particle, not a wave. That might seem far removed from Schrödinger's cat – a cat, after all, is a very different beast from an electron. But subsequent experiments have pushed the quantum particle to ever-larger sizes.

We have carried out the quantum double slit experiment with photons, electrons, atoms, and even 60-atom fullerene molecules. The weird interference effect has never disappeared –

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 34

unless we tried to look at which slit the particle went through. Plans are afoot to do it with much larger objects: a virus, and maybe something a million times bigger than the fullerene molecule. Apart from the difficulties of building the experiment, there is no fundamental reason to stop there: there is no cause to suggest why a real cat shouldn't behave in the same way as an electron, given the right circumstances and a cat-flap-sized double slit.

Except, of course, that it is easy to see a real cat, and thus determine which cat-flap it went through. In Schrödinger's thought experiment, the box has to remain closed so that no one can see the cat, there is no measurement performed, and the superposition remains intact. This leads us to a difficult question, one that Bohr always evaded. What constitutes a measurement? With Schrödinger's cat, is it when the box opens? When light photons bounce off the cat relaying to us the information that allows us to tell whether the cat is dead or alive? Or is it when those photons enter our eyes? Or when our conscious minds register the state of the cat? Bohr's answer to this conundrum was, essentially, that physicists just know when they have made a measurement. Modern versions of the Schrödinger's cat experiment, however, are shedding much more light on the process – and explaining why a cat can't really be dead and alive at once.

Don't look now

The boundary between the 'classical' world that we inhabit and the quantum world of the atoms comes down to the de Broglie waves that brought this whole story into existence. The de Broglie wavelength of a body, which depends on its momentum, gives a measure of the scale at which it will manifest as a quantum wave.

In the double slit experiment, the fullerene molecule has a de Broglie wavelength of around 10^{-12} metres, or a thousand billionth of a metre. The gap between the slits is around half a million times bigger than that; bigger, but not too different in scale. This means the system is suited to exposing wave behaviour.

This is still in line with Bohr's claim that the choice of measurement apparatus decides which characteristics will manifest, but it does throw out two explanations for why a cat or a person can't – unlike the fullerene molecule – seem to be in two places at once.

The first reason is practical. Walking along a wall at a couple of miles per hour, for example, Schrödinger's cat would have a wavelength of around 10^{-28} metres. Its quantum, wave-like behaviour would only be exposed by a measuring device of a similar scale. Since we have never created such a device, we cannot perceive quantum behaviour. Everyday life is, according to Bohr's scheme, an experimental situation that will always manifest the particle-like nature of everything around us.

The second reason that we are 'classical' is that we are emitting radiation. Anything that has a temperature above absolute zero, –273 degrees Celsius, emits photons, packets of energy that carry away heat. Experiments have shown that this radiation can be used to find the location of the object, effectively revealing which slit it passed through. In other words, at a temperature above absolute zero, you can't close the box on Schrödinger's cat, invalidating the premise of the thought experiment whenever you translate it into the real world.

These experiments were carried out by firing fullerene molecules at a double slit. The hotter the fullerene molecule was as it approached the slits, the more blurred the interference pattern. The hot molecule emits photons, and the energies of the emitted photons are determined by the temperature. Higher temperature essentially gives higher energy, which translates, in de Broglie's terms, to a shorter wavelength. And the shorter the wavelength of the emitted radiation, the easier it is to infer the emitting molecule's position. In other words, a hot body seems to give away more information about which slit it might go through.

The same thing happens if the fullerene molecules collide with air molecules on the way to the slits. Normally the

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 36

experiments are done in high vacuum, but if the vacuum is not so good, and the position of the fullerene can be inferred by watching what it does to air molecules, the interference pattern fades away. Again, as it becomes possible to infer which slit the molecule goes through, its ability to go through both at the same time begins to disappear. In a partial vacuum, the fullerene behaves as if someone had left the box half-open on Schrödinger's cat, forcing it to be alive or dead, but not both.

COMPUTING WITH CATS

The idea behind a quantum computer is to use the Schrödinger's cat phenomenon to perform computations on a massive scale. Familiar computers use the charge state of a capacitor to represent a number in binary: 0 or 1. Quantum computers, on the other hand, use the state of an atom. If it is in its normal state it is 0. If it is given a little extra energy, it is 1. But, being a quantum object, the atom can be in a superposition of 0 and 1 at the same time.

Using another quantum phenomenon called 'entanglement' to string together lots of atoms in superposition allows quantum computing researchers to create a string of undetermined numbers that, when put through a series of steps, perform computations on all possible numbers at once. Quantum computing is a way of doing 'parallel' computations on an unprecedented scale. In theory, an entangled string of just 250 atoms, each in a Schrödinger's cat superposition state, can encode more numbers than there are atoms in the universe. The potential is huge. No wonder governments are seeking to protect their national security ciphers from the developers of the first quantum computer.

Except for the problem of decoherence, that is. The nature of entanglement and superposition make the atoms especially vulnerable to losing information, and when they do, the computation falls apart. If researchers could get more of a handle on decoherence, and why we never see alive-and-dead cats, they might be able to usher in a revolution in computing.

So, information does not have to enter a conscious mind to constitute a measurement: it just has to leak away from the system under scrutiny. It appears that a flow of information about the health of Schrödinger's cat is enough to force it into one of the two states available. Where humans and cats are concerned, that information leaks away because our bodies interact with our environment in myriad ways, radiating heat and knocking air molecules about. Information about the position of our bodies is available, which means we can't be in two places at once. This spilling of information is known to scientists as 'decoherence'. Decoherence is not a trivial issue: it might just show us the very nature of the universe.

> *'Anyone who is not shocked by quantum theory has not understood it.'*
>
> NIEL BOHR

Information and reality

The physicists looking into the enigma of Schrödinger's cat are now wondering whether it points to the notion that information is the most fundamental element of reality. Quantum theory, in the form of Schrödinger's unfortunate cat, suggests that the universe can be described as one giant information processing machine. And this leads to potential applications too. The role of information in quantum theory has led us to one of our most ambitious technological projects: the quest to build a super-powerful processor called a quantum computer (see Box: Computing with cats).

However powerful it turns out to be, the quantum computer is unlikely to be able to help us comprehend how a cat really can be alive and dead at the same time. The idea that this is part of the nature of physical reality remains truly outrageous to the human mind. Wolfgang Pauli, who didn't give up on physics, and became one of the most brilliant physicists in the history of science, was right. It's too complicated to grasp. As Niels Bohr once said, 'Anyone who is not shocked by quantum theory has not understood it.'

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 38

WHY DOES AN APPLE FALL?

Gravity, mass and the enigma of relativity

Because of gravity, of course. Everybody knows that. However, what is the fundamental nature of gravity? That is a much harder to question to answer, despite the fact that gravity is the first of nature's fundamental forces to penetrate your consciousness.

Here's an experiment you can try at home. You'll need a six-month-old baby (you could borrow one). Tie a piece of fishing line to one of the baby's toys – a rattle, say. Now suspend it from the ceiling at a height where it will rest lightly on a chair with the line taut and invisible. Get the baby to watch as you whip the chair away. Keep your eyes on the baby's: when, for no obvious reason, the rattle doesn't fall to the ground, the baby will stare at it for much longer than is normal.

This, according to psychologists, is how babies express astonishment. It seems that we know from a surprisingly young age that things are meant to fall downwards when unsupported, and we are mystified if they don't. No wonder the levitation tricks of the Victorian illusionists entranced an entire generation. When things cheat gravity, our very core takes delighted offence.

Gravity, you see, is a tyrant. It cannot be cheated. We cannot, as we can with an electric or magnetic field, block it out. Neither can we counter it with an opposing force – there seems to be nothing in physics that equips us with antigravity. The rule of

gravity is so central to human experience that we have become, essentially, oblivious to gravity's presence. It is only in its absence – or, rather, its apparent absence – that we remember it is always there.

Perhaps that is why the earliest ventures into science largely ignored gravity. As we understand it now, one type of action governs the fall of a tripping human, the arc of an arrow's flight and the motion of the planets, but Aristotle's textbook *Physics* makes no mention of any universal force orchestrating the universe. He did suggest that objects did not fall off the earth because of the earth's 'heaviness', but his reasoning was askew. He suggested that the strength of the earth's pull depended on how big an object was and what it was made of.

In Aristotle's view, heavy objects fall more quickly than light objects. That is because of the Greek obsession with the elements: Earth, Air, Fire and Water; most of the heavy objects Aristotle knew about were made from materials found in the earth, and the strong pull, he said, was because they were compelled to return there. Our understanding didn't really move on from this flawed idea for almost 2,000 years. Eventually, though the Italian scientist Galileo Galilei established that Aristotle was wrong: heavy objects are not more strongly attracted by the earth. As long as air resistance is not a factor, a heavy and a light object will fall at the same rate.

As easy as falling

Sadly, the romantic stories about Galileo's proof of this – by dropping cannonballs from the leaning tower of Pisa – are not true (the myth was started by Galileo's student Vincenzo Viviani), but it has nevertheless been proved in an even more spectacular fashion. In 1971, Apollo 15 astronaut David Scott paid tribute to the discovery's profound consequences by dropping a hammer and a falcon feather onto the moon's surface. 'One of the reasons we got here today, was because of a gentleman called Galileo,' Scott said as he let them fall. The hammer and the feather landed, of course, at the same time.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 40

Scott's appraisal was almost correct: astonishingly, it really didn't take much more than Galileo's 17th-century insights to get us to the moon. The gaps were filled by a man born just one year after Galileo died: Isaac Newton. Unimpressive as he was at birth – his mother said he could be 'put in a quart mug' – Newton took just a couple of dozen years to gather all the information it would take to plot a course for the Apollo astronauts four centuries later. And here, of course, is where the apple comes in.

Unlike the stories of Galileo's experiments on the leaning tower of Pisa, accounts of Newton's gravitational epiphany at the sight of a falling apple are almost certainly true. It was late summer, 1666, and Newton was sat in his garden at Woolsthorpe Manor in Lincolnshire. The apple tree is still there, and still bearing fruit every autumn.

An apple falls because it has a property called mass, and so does the earth. Newton's great leap forward was to spell out how everything with mass attracts everything else with mass. His universal law of gravitation, constructed at the tender age of twenty-three, said that the attractive force is dependent on those two masses, the distance between them, and a constant known as G.

Actually, physicists are often over-familiar with the gravitational constant and call it 'Big G' to distinguish it from (little) g, the acceleration due to the earth's gravitational pull. However, despite the familiarity, G is actually the least well defined of all the fundamental constants.

The size of G, like that of all the other fundamental constants, is known not through some theoretical argument, but through measurement. The English physicist Henry Cavendish was the first to measure it, in 1798, by analysing the gravitational attraction between two known masses that were a known distance apart. His answer for G was 6.754 x 10–11 metres cubed per kilogram per second squared. Today, G is officially 6.67428 x 10–11 m3/kg/s2. The uncertainty on this measurement is about one part in 10,000. Compare that to the precision with which we know the other

fundamental numbers, such as the Planck constant used in quantum theory: that is known to 2.5 parts in 100 million.

There are two reasons G is so difficult to measure accurately. The first is that it is impossible to screen out gravitational fields using any known physics. That means any measurements have to take into account the influence of any and all objects in the vicinity. This makes the measurements unreasonably sensitive to external influence; there are stories of researchers having to recalibrate their apparatus after someone two laboratories away has moved a large pile of books into their office. For this reason, gravity measurements have to be done in isolated laboratories using extraordinarily sensitive instruments.

The second difficulty with measuring the gravitational constant is the fact that gravity is the weakest of the fundamental forces. When that apple falls to the ground, it does so with relatively little acceleration, despite the fact that the mass of the entire planet is tugging it downward.

If you're not convinced that gravity is weak – maybe you've done a parachute jump or been on a rollercoaster and experienced a terrifying acceleration – think about the magnets sitting happily on the door of your refrigerator. The mass of the entire planet is working to pull them towards the ground too – and yet a button-sized dot of magnetized iron can resist the planet's pull. Magnetism results from the electromagnetic interaction between charged particles inside a magnet. And that force is around 1042 – that is, around 1 million trillion trillion trillion – times larger than the gravitational force between them. So gravity is weak: G is astonishingly small. But why? Though the weakness of gravity is one of the central mysteries of physics, we do have some ideas that might account for it. The best is that gravity 'leaks' into or out of our universe.

Leaks from another world

Various branches of modern physics suggest that there are many more dimensions of space than the three (up and down, side to

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 42

side and backward and forward) that we are familiar with. One of the consequences of this is that certain forces can become 'diluted' through spreading into these extra dimensions. If the gravitational force is weak, that may be because it is spread more thinly than the others.

The 'extra' dimensions are thought to be 'compactified' – rolled up, essentially – so small we don't experience them in day-to-day life. It's just a theory at the moment, but a few researchers are trying to find evidence for this. One route is through examinations of the way the gravitational attraction between two objects changes with the distance between them.

'One of the reasons we got here today, was because of a gentleman called Galileo,'
DAVID SCOTT

Newton showed that gravity follows an 'inverse square law'. That means that the gravitational force one object exerts on another decreases in proportion to the square of the distance between them. Separate two objects by a metre, and measure the gravitational force. Then separate them by a further two metres and measure the attraction again. It will be nine times weaker because they are three times further away.

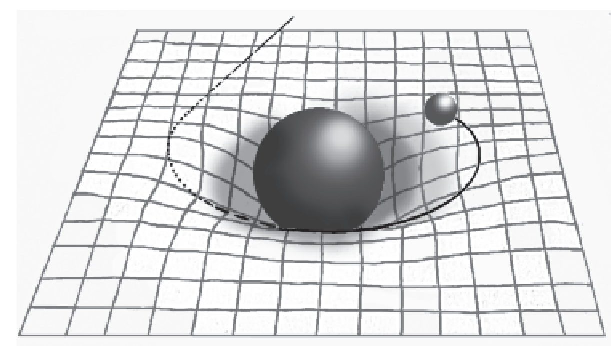
The hidden dimensions enter our world at submillimetre scales. If gravity behaves differently from normal on these very small scales – if the inverse square law doesn't hold when masses are separated by only a few thousandths of a millimetre – that may be because these dimensions are interfering with things. Spot some disturbance here, then, and we might have evidence to support our most daring theories.

This is why physicists are carrying out the most exquisite experiments to probe gravity at microscopic scales. So far, however, they have found no evidence of violations of the inverse square law. That's a great shame, because one of the roles of these advanced, multidimensional theories is to improve our best theory of gravity, Einstein's relativity.

Gravity is relative

Einstein's theory of relativity cast space and time as a four-dimensional fabric and said the presence of mass or energy distorted this fabric. Where Newton had declared that bodies in motion will move in a straight line unless acted on by a force, Einstein added a twist. Yes, they moved in a straight line through space, but would have to follow any distortions in that space.

The distortion that the sun's mass creates, for instance, means that a nearby planet in motion will be pulled into a curved trajectory. Balance the masses and the speed of motion, and you have an orbit. Hence, in Einstein's view, gravity is a kind of illusion. Though it looks like a force that acts across space and time, it is actually more like topographical features – hills and valleys – added to the landscape, features that make it hard to travel in certain directions, and easier to travel in others



THE LARGER BODY'S GRAVITATIONAL FILED 'BENDS' SPACE TRAPPING THE SMALLER BODY IN AN ORBIT

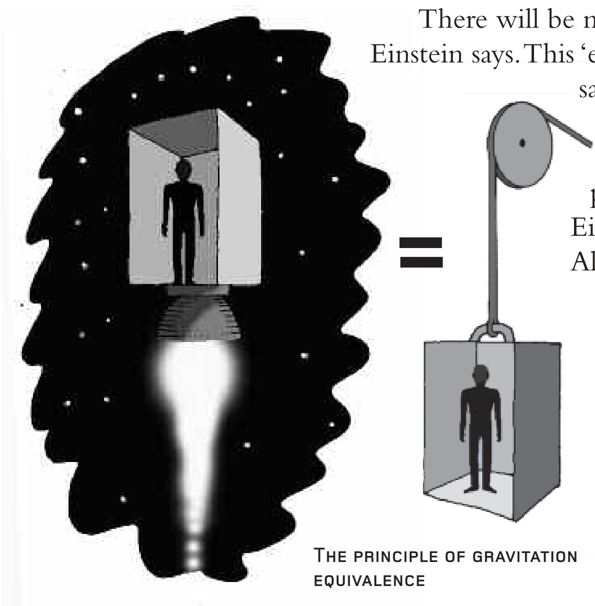
Neat though this is, and supported by numerous experimental findings, we know it is not the final answer. In a way, Einstein has only given us a clever description of *how* gravity works. The *why* is still wide open. There is hope, though. Relativity, in its current form, is not compatible with quantum theory. We will have to wait for some future 'quantum gravity' theory to unite the two. And that theory, presumably, will give us the why of gravity, just as we have recently got to grips with the why of mass.

So far, we have been blithely talking about mass, while avoiding the obvious question. What does it mean that something has mass?

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 44

Physicists categorize mass in two distinct ways. One is 'gravitational mass', which is what produces and responds to gravitational fields. This is what makes the apple fall. The other is 'inertial mass', which is a measure of how hard it is to move something out of its current state of motion or rest. When you try to push a broken-down car, its inertial mass stands against you.

As far as we know, inertial and gravitational mass are entirely equivalent. Imagine standing on earth in a sealed box like a stationary elevator. You feel the push of the floor as your gravitational mass responds to the influence of gravity. Now imagine taking that elevator box into space, away from gravitational fields, and sticking a rocket engine on it that accelerates it at 9.81 metres per second per second, the acceleration due to gravity at the earth's surface.



THE PRINCIPLE OF GRAVITATION EQUIVALENCE

There will be no difference in what you feel, Einstein says. This 'equivalence principle', which says there is no distinction between your gravitational mass and your inertial mass, is part of the bedrock of Einstein's general relativity. Although we don't have a definitive proof that it is absolutely correct, experiments have shown this is certainly true to at least one part in 10^{12}. A decade before he created general relativity, though, Einstein asked another question about mass. In 1905, his 'miracle year' when he also published his special theory of relativity, Einstein came up with an interesting line of enquiry. He asked, in a landmark paper, whether the inertia of a body depends upon its energy content.

The energy of weight

This was the origin of the world's most famous equation: $E = mc^2$ (see *Why does $E = mc^2$?*). Energy and mass, in Einstein's view, could be interchanged. It took almost a century, but we now know through this equation that energy is indeed the root of mass. Take that apple, for instance. Its mass resides in its constituent components. Going down in scale, these are molecules, which are composed of atoms, which are composed of electrons, protons and neutrons.

The origin of the mass of the electron (which is only one-thousandth the mass of the proton and neutron) remains a mystery. But physicists are at least getting to grips with the mass of protons and neutrons. These particles are each composed of three particles called quarks. However, the masses of the quarks account for only around 1 per cent of the proton or neutron mass. The rest comes from a shadowy, quantum world of energy-stealing 'virtual particles'.

Down at the quantum scale, the rules are very different from those that we encounter in our day-to-day lives. Here, a phenomenon called the 'Heisenberg uncertainty principle' holds sway, and issues strange declarations. One is that nothing has a definite amount of energy, even when that energy is zero. Instead the energy fluctuates around zero, allowing seemingly empty space – something physicists call the 'vacuum' – to fizz with appearing and disappearing particles.

These particles appear in pairs: a particle and its antiparticle, spontaneously created as the energy of the vacuum of empty space fluctuates around zero. According to a Nobel Prize-winning branch of physics called 'quantum chromodynamics' (QCD), the particles can appear with various amounts of energy, giving a spectrum of characteristics. Sometimes they take a form where they are known to physicists as 'gluons'. Gluons create a force known as the strong nuclear force, which holds quarks together to create a proton or a neutron. And it is gluons – or rather their energy – that give the apple most of its mass. Working out exactly

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 46

how much mass comes from the energy of all these virtual particles has not been easy, involving crunching combinations of around 10,000 trillion numbers. When the results came out, though, they were within a couple of per cent of the experimentally recorded masses of these particles.

The energy associated with the gluons, converted via Einstein's $E = mc^2$ formula, accounts for almost all the mass in a proton or neutron. There is a little missing: the mysterious electron mass, and a contribution from some more virtual particles, such as pairs of virtual quarks and antiquarks, and the Higgs boson (see *What is the God Particle?*). Essentially, though, the mass of the apple – and of the earth – is a manifestation of the energy contained in the vacuum of empty space.

The success of quantum chromodynamics in establishing the origin of mass has given physicists hope that similar ideas will eventually lead us to the final *why* of gravity: the graviton. The electric and magnetic forces are manifest through atoms exchanging of packets of energy called photons. The strong nuclear force comes via gluons, as we have seen. The weak nuclear force is known to result from the exchange of energy-laden particles known as the W and Z bosons. All of these have been seen in experiments. Gravity is thought to rely on the exchange of particles known as 'gravitons'. These, however, remain hypothetical. Despite all our advances in understanding, we still haven't seen a graviton.

That is not our only remaining problem with gravity, however – a much more embarrassing and basic issue remains unsolved. Bizarre as it may seem, although we have worked out the origin of mass using the most ingenious minds, the biggest computers and the greatest theories of physics, we still don't have a good way to measure the very thing that gravity acts upon: mass. Every other standard unit of measurement has a precise, atomic foundation. The second is based on a certain number of oscillations of a caesium atom. The metre is the distance light travels in a particular fraction of that second. The kilogram,

though, is the mass of a lump of metal kept locked inside a Paris vault.

The changing kilo

It's not any old vault, of course: it is contained within the hallowed walls of the International Bureau of Weights and Measures (BIPM) near Paris. And it's not any old metal, either: it is a cylinder of platinum, chosen as the most stable, incorruptible material available. The mass of this platinum cylinder is the kilogram against which all other kilograms are calibrated.

The problem is, its mass is changing. Metrologists have made dozens of copies, and the original no longer weighs the same. There is about 100 micrograms of difference, roughly equivalent to the mass of a couple of grains of salt. Researchers are planning ways to bring the kilogram into line with other standards, by using atomic measurements. One hope is to create a polished sphere of silicon containing a determinable number of atoms. The kilogram will then be defined as the mass of a certain number of silicon atoms.

Another possibility is to use something called a Watt balance to measure mass in terms of energy. Einstein told us that mass and energy are interchangeable; the Watt balance would invoke this by measuring mass against the energy contained in a carefully configured electromagnetic field. Until these plans come to fruition, though, we are stuck with plugging slightly inaccurate numbers into Newton's formula.

Gravity is everything to us – it pulled particles together to create the earth, it holds us in orbit around our life-giving sun, it creates the tides that allowed that life to form and move on to land. And now we return the favour and use our gravity-given minds to make extraordinary discoveries about the very nature of this attraction. At the same time, though, we have only primitive tools to get the measure of it. While we can talk about gluons inside the nuclear structure of the atoms within the apple, we cannot be precise about how much the apple weighs. The essence of gravity remains deliciously difficult to tame.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 48

ARE SOLIDS REALLY SOLID?

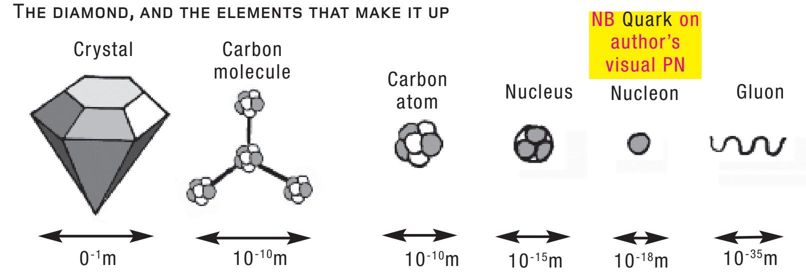
Atoms, quarks and solids that slip through your fingers

If all the world was made of gas, we could not exist. The way our bodies are organized, the way information is stored in the structure of DNA, the way our brains process and hold information, all requires that atoms are fixed in place, not floating freely around. Life, at least life as we know it, requires solidity. But what is a solid?

A gas is a collection of atoms or molecules that have no or extremely weak bonds between them. A liquid has weak bonds between the particles, allowing them to slip over and past each other. A solid, though, has its particles held together by strong electrostatic bonds. But that doesn't make a solid solid. Hold your hand up in front of your face. It looks solid enough, doesn't it? But to neutrinos, the tiny subatomic particles that flood the universe, your body is far from solid. Every second, trillions of neutrinos pass right through you, failing to interact with a single atom of your body. Scientific progress has made it clear that most of our solid matter is empty. We have even devised solid materials with the ghostly power to pass through each other. Experimental science is teaching us that the concept of 'solid' is a slippery one at best.

Our brains, another mass of solid atomic matter, have been able to probe this a level even deeper than our experiments. Though there is no certainty here yet, our best understanding leads us to a remarkable conclusion: that there is no such thing as a solid.

THE DIAMOND, AND THE ELEMENTS THAT MAKE IT UP



Every piece of matter is, essentially, the result of a random fluctuation in the energy of space and time. Solidity is, at its root, an illusion.

To explore this, let's start with a familiar solid. Something dependable, something robust. Diamond seems a good solid to test. It is the hardest naturally occurring material, and used as a tool to cut through the toughest of metals. How solid is diamond? It is diamond's molecular structure that makes it particularly hard. Its carbon atoms are bonded in a rigid tetrahedral arrangement, sitting about 10^{-10} metres apart from each other. Since it is the outermost electrons in the atom that form these bonds, it will come as no surprise to hear that this is roughly the size of the atom. But that doesn't make it truly solid. It is time for us to explore the strange world of atomic structure.

The first scientist to look into this question is generally considered to be Democritus. He was actually a Greek philosopher rather than a scientist, but made a scientific conjecture about the nature of matter. All matter, he suggested, can be split so far, but no further. At the most fundamental level was the concept of *atomos*, from which we get our word atom. In Democritus's view, *atomos* were the particles that could not be split, destroyed or changed in any way.

And, until the earliest moments of the Industrial Revolution, that was essentially that. The age of the telescope came, and we learned to probe the heavens, but we made no progress in getting to the root of matter. That's because we needed tools that could influence matter on the atomic scale.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 50

Inside the atom

It was the English schoolteacher John Dalton who kicked off the investigation of the atom. Towards the end of the 18th century, Dalton proposed that any single element was an assembly of identical atoms. These all had the same properties. Chemical reactions, he suggested, joined two different kinds of atoms together to form a chemical molecule. Dalton backed up his ideas with chemical experiments that determined the ratio of elements within certain substances, such as carbon dioxide: one part carbon to two parts oxygen.

> *'It was as if you had fired a fifteen-inch shell at a piece of tissue paper and it had come back and hit you.'*
>
> ERNEST RUTHERFORD

The concept of atoms lent itself to the processes of the Industrial Revolution, enabling the pioneers of thermodynamics to work out the gas pressures and heat transfer rates that powered the rise of the machine. But we were no wiser about whether it might be possible to get inside an atom. In the age of the British Empire, the steam train and massive industrialization, the science of the atom had hardly moved on from the Greek idea of an indivisible substance.

Three near-simultaneous developments changed that. The investigations of English physicist Joseph J. Thomson revealed the existence of particles, which he called 'corpuscles', that were negatively charged and were 2,000 times lighter than even the lightest of the atoms. With this discovery, we had at last found something – we now call it the electron – that was smaller than an atom.

By 1904 Thomson was suggesting that atoms were composed of positive and negative parts mixed together to give a 'plum pudding' kind of structure. Around the same time, in Paris, Pierre and Marie Curie and Henri Becquerel discovered radioactivity. Their subsequent investigations showed that at least some of the activity resulted from the emission of charged particles from atoms. Back in England, meanwhile, the brash New Zealander Ernest Rutherford had arrived. In just a few decades of research, Rutherford was to make the greatest inroads into the atom for thousands of years.

Perhaps the most significant discovery was the revelation that Thomson's 'plum pudding' model of the atom was entirely wrong. Rutherford fired positively-charged alpha radiation particles – helium atoms stripped of their electrons – at a thin piece of gold foil. Almost all of the alpha particles passed straight through. Some, however, were strongly deflected. A few even bounced back at the emitter. This shocked Rutherford. 'It was as if you had fired a fifteen-inch shell at a piece of tissue paper and it had come back and hit you,' he later wrote.

Nuclear bombshell

To Rutherford, there was only one interpretation of this extraordinary result. A few of the positively charged helium atoms had happened to be fired directly at a concentration of positive charge, and been strongly repelled. Most of the volume of the atom was empty space. But at the centre lay all the positive charge – and almost all the mass. Rutherford had discovered the atomic nucleus.

The emptiness of an atom is hard to grasp, and provides our first clue to the illusion of solidity. The nucleus in the atom,

IS THE ELECTRON SACRED?

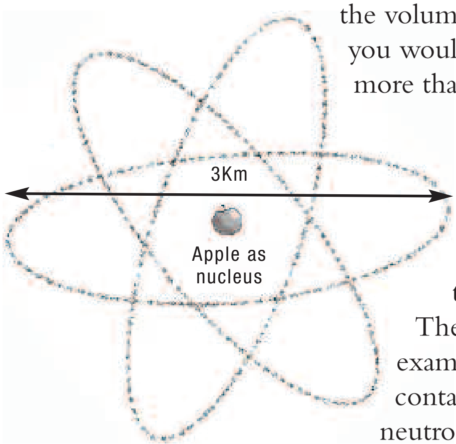
The history of physics has dealt a series of blows to our hopes of finding the fundamental particles of matter. The discovery that an atom could be split, and the subsequent discovery of the nucleus and all its constituent parts, has taken us deeper and deeper into the puzzle of solidity. The one thing that was thought to be fundamental is the atom's negative charge, the electron. But even here, there are now doubts.

In 1998, a trio of physicists won the Nobel Prize in physics for demonstrating that the electron could behave as if it was split into three parts. It is still not clear whether the electron actually splits, but the electron is something of a mystery anyway. We have no idea, for instance, where its mass comes from. Since we have found that protons and neutrons can be split, and their mass ultimately comes from quantum fluctuations, when it comes to the solidity of the electron, all bets are off.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 52

The nucleus in the atom, Rutherford said, is 'like a gnat in the Albert Hall'

Rutherford said, is 'like a gnat in the Albert Hall'. Others around him called it 'the fly in the cathedral'. Either way, it's a monstrous emptiness. If the nucleus was the size of a small apple, the edge of the atom, defined by the outer orbit of its negatively charged electrons, would be 3 kilometres (2.5 miles) in diameter. Each electron, meanwhile, would be smaller than the full stop at the end of this sentence. We can look at the emptiness another way. If you could remove the empty space in atoms, and pack hydrogen nuclei into the volume of a penny with no space between them, you would have a penny-sized object that weighed more than 30 million tons.



THE EMPTINESS OF THE ATOM

Inside the nucleus

As the lightest element, hydrogen has the simplest possible nucleus: a single positive nuclear charge, or proton. But, generally, there is more to nuclei than just the proton. The carbon atoms we have been examining, for example, have a much more complex nucleus, containing half a dozen uncharged particles called neutrons. All atoms (apart from hydrogen) contain neutrons. The neutron, which is very slightly heavier than the proton, was discovered by James Chadwick at the University of Liverpool in the early 1930s. Carbon has a nucleus composed of six protons and, depending upon the exact 'isotope' we are dealing with, six, seven or eight neutrons.

So, is there any solidity here? Rutherford found the proton to be around 10^{-15} metres in diameter. The neutron is of approximately the same size. And atomic nuclei don't mirror the emptiness of the atom. The carbon nucleus is no bigger than one would expect if the particles within it were packed tightly together. Larger nuclei make the tight packing of the nucleus even clearer. A uranium nucleus, which contains 238 particles, is only 14 proton-widths in diameter – it is rather like a basketball stuffed with 238 ping-pong balls.

With this discovery, physicists had a notion of solidity at the core of matter. But only for a while: things soon got slippery again, taking us on a downward spiral that, today, tells us there is nothing solid in the entire universe. The problem is that, being packed with positive charges, the nucleus should not hold together. The protons in a carbon nucleus should, by rights, repel each other.

That means another force must be at work. Physicists call this the 'strong' nuclear force simply because it has to be strong enough to overcome the repulsive electromagnetic force. To investigate the strong force required physicists to delve into the characteristics of the proton and neutron, or nucleon, as they are collectively known. What they discovered was that the nucleons were not fundamental, indivisible particles, but composed of three 'quarks'.

Quarks at heart

The name 'quark' was chosen by physicist Murray Gell-Mann in 1964, who picked it out after reading the phrase 'Three quarks for Muster Mark' in James Joyce's *Finnegan's Wake*. The quark started out as a hypothetical particle, whose existence was also suggested, independently, by the Russian-American physicist George Zweig (who wanted to call it an 'ace'). Both men's guess turned out to be a good one – though it took a good while to prove it.

Physicists can only see matter on this scale by smashing together subatomic particles in accelerators. The collisions create smaller particles, whose fleeting existence has to be inferred by the trails left behind in detectors that line the walls of the accelerator at the collision site. The first quarks were identified from collisions at the Stanford Linear Accelerator Centre (SLAC) in 1968. Two more decades passed before all the hypothesized quark particles had been seen. But we now know that quarks come in six 'flavours', exotically named: strange, charm, top, bottom, and the much more common up and down.

Protons are composed of two up quarks and one down quark; neutrons are two down quarks and one up quark. But it is the top quark that may be the undoing of solidity. The top quark is

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 54

unaccountably heavy. It weighs almost the same as a gold atom, which is why it took until 1995 for our particle accelerators to be able to make one. Particle accelerators are governed by $E = mc^2$, and it takes a great deal of energy to make so much mass.

A gold atom contains 79 protons and 118 neutrons. That is a total of nearly 600 up and down quarks. How can just one top quark weigh nearly the same? Something in the nature of quarks, and how they come together, suggestes that there is a mystery in the nature of mass. A theory called quantum chromodynamics (QCD) makes that clear. It has shown that the up and down quarks that make up protons and neutrons account for only 1 per cent of the mass of these particles. The rest is, as provided by $E = mc^2$, held in the energy that binds the quarks together. This is the 'strong' nuclear force.

Sensing the energy of emptiness

According to QCD, the strong force has its roots in the uncertainty principle of quantum mechanics (see *Is Everything Ultimately Random?*). This principle says that nothing that can be measured actually has a precisely defined value. That even applies to empty space: it can't have exactly zero energy. As a result, empty space has a fluctuating but finite amount of energy.

This fluctuating energy manifests as particles called gluons, and it is gluons that create the strong force that binds the quarks. So when you hold a diamond in your hand, you feel its weight. But what you sense as the mass of the diamond is actually the result of a shifting, shimmering energy field that creates the weight of the quarks that make up the protons and neutrons in the nucleus of each carbon atom. In a sense, that diamond, that most solid of objects, doesn't have a permanent existence at all. As it rests on your hand, all that is happening is that a continuum of energy fluctuations are manifesting as solidity.

Slippery solids

Perhaps we shouldn't be surprised that the rules of solidity are turning out to be flexible. Solids, after all, are only solid under certain

conditions. Heat up an ice cube, and it will create a pool of water. The molecules haven't changed their essential nature; it is simply that the environmental conditions have altered the strength of the bonds between them. The same is true when we heat the water and it turns into steam. Now the bonds between the molecules have disappeared – but still the molecules themselves haven't changed.

We can create a new kind of matter, at the other end of the temperature scale, too. As we cool some kinds of materials down, we can create a new kind of matter. To solid, liquid and gas, we can add the phase known as the 'Bose-Einstein condensate'. The BEC is the result of a radical transformation that only happens at extremely low temperatures. Temperature is, in essence, a measure of how much energy an object has to 'jiggle about'. At very low temperatures, a material is stripped of all energy, and so hardly moves at all. But quantum theory dictates that the more precisely you pin down an object's momentum – in this case to near-zero – the bigger the uncertainty in its position. So every particle in the BEC has an uncertain position. In effect, all the particles overlap each other, merging into one big quantum object, like a giant atom.

In this state, all kinds of strange behaviours arise. When niobium metal turns into a BEC, the quantum laws turn it into a 'superconductor' that carries electrical current without any of the resistance associated with currents in normal metals. When helium atoms form a BEC, for example similar thing happens: stir a cup of this 'superfluid' helium, and the swirl goes on swirling for ever. Even more bizarrely, superfluid helium can defy gravity, flowing up the sides of a container. Turn helium into a solid, where its atoms are held together in a crystal, and the weirdness gets worse.

Not that it's easy to make solid helium. To get it to a liquid requires cooling it to within 4 degrees of absolute zero. To turn that liquid into a solid requires crushing the atoms together: the liquid has to be cooled to within 1 degree of absolute zero and compressed with 25 times normal atmospheric pressure. Once you're there, though, you can see the strangest solid in the universe.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 56

The bonds between the atoms in solid helium are extremely weak. So weak, in fact, that atoms can break off. This leaves what is known as a 'vacancy' in the crystal. Physicists have long known that these vacancies can be treated like particles in their own right. They are really like an atom with slightly different properties. They affect the way a material conducts electricity, for example; it is only because of vacancies that semiconductors have the properties they do. The entire multi-billion dollar business of electronics relies on the properties of vacancies.

In an ultra-cold helium crystal, the laws of quantum mechanics lock all the vacancies in the structure together to form a vacancy-based BEC. With the atoms locked together too, the helium crystal becomes two 'supersolids'. And, if you get the experimental conditions right, they can pass right through each other. In theory, any solid crystal will behave in this way under the right conditions.

It might not even require the formation of vacancies: in some materials it ought to be possible to make all the freed atoms lock together and move around the crystal as one, meaning that the solid will pass through itself. It's not unlike the strange conjuring tricks where two solid rings are made to pass through each other, lock together and then, with a flourish of the magician's hand, come apart again. In this case, though, it is the solidity that is the illusion.

Look at your hand again. It is mostly made of nothing. The crystal structures of the proteins leave enormous gaps between the tiny atoms. The atoms themselves are almost entirely devoid of matter. Where there is matter – in the atomic nucleus – most of its mass is derived from quantum fluctuations in the energy of empty space. The solidity of that hand in front of your face is perhaps the most convincing illusion you will ever experience.

WHY IS THERE NO SUCH THING AS A FREE LUNCH?

Energy, entropy and the search for perpetual motion

The exact origins of the phrase 'no such thing as a free lunch' are unclear, but most sources say it began life as the pithiest summary of economics. It appeared in Pierre Dos Utt's 1949 monograph 'TANSTAAFL: a Plan for a New Economic World Order', where Dos Utt tells of a king seeking economic advice. His advisers, looking for ever-simpler ways to get their message across, conclude with the now-classic version of the phrase: 'There ain't no such thing as a free lunch.'

It is doubtful this would have been enough to motivate economists to usher in a new world order, and the physicists of the time would have certainly been unimpressed. The idea of something for nothing had long been a goal of inventors trying to get a free lunch by coming up with 'perpetual motion machines' that would do work without the need for any external power. Physicists had long been telling them this was impossible.

There is no such thing as a free lunch because you simply can't get something for nothing: someone, somewhere always has to pay. Physicists have enshrined this principle as a fundamental law of physics. So you need to think hard before you start looking for a free lunch, because you are battling against the way the universe runs. Perhaps the great artist, visionary and inventor Leonardo da Vinci put it best. He took a keen interest in perpetual motion, investigating designs, and coming up with a few of his

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 58

own. But he was sceptical about them all: one of his notebooks contains a detailed analysis of a popular kind of machine, showing why and how it could not work. 'O you researchers of perpetual motion,' Leonardo wrote, 'how many harebrained ideas have you created in this search. You may as well join the alchemists.'

'O you researchers of perpetual motion, how many harebrained ideas have you created in this search. You may as well join the alchemists.'

LEONARDO DA VINCI

There are two kinds of perpetual motion machines. The first supplies an endless output of work despite the fact that there is no input of fuel or any other form of energy. The second converts heat to mechanical work with perfect efficiency. Both, it should be made clear, are wishful thinking – and physics tells us why.

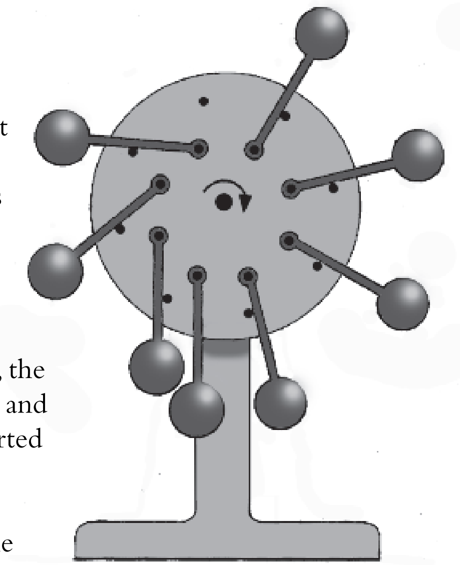
Something for nothing

As with alchemy, the search for perpetual motion engaged some of the finest minds that have graced the Earth. The dream has been around since at least AD 624, when the Indian mathematician and astronomer Brahmagupta described a wheel whose hollow spokes could be filled with mercury. The mercury would shift weight around the wheel as it rotated. As a result, Brahmagupta wrote, 'the wheel rotates automatically for ever'.

The idea was repeated numerous times. In 1235 Villard de Honnecourt, a French artist and inventor, produced his own version. De Honnecourt was no fool: he drew the first known plans for a mechanical escapement mechanism that would keep time. But de Honnecourt's 'overbalanced wheel' still doesn't work. Here, a series of hinged weights are attached around the circumference of a wheel, their motion limited by pins. As the wheel turns, an imbalance in the distribution of weights causes the wheel to turn. As it turns, the elevated weights drop onto their pins, and the transfer of weight keeps the wheel turning.

The fact that the perpetually rotating wheel is a running theme in the search for perpetual motion can only mean that very few people tried to build these kinds of machines. Build one

and you soon learn that they just don't work. Take de Honnecourt's overbalanced wheel, for example. What is needed for this to carry on for ever is for the uppermost rod to flip over as it reaches the top of the wheel, maintaining the imbalance. Unfortunately, this doesn't happen: the weight distribution is such that it doesn't quite flip. After one revolution, the weights return to their initial position, and everything is back exactly where it started – including the stationary wheel.



THE OVERBALANCED WHEEL

To be fair to de Honnecourt, the reason for this was not clear until well after his time. The problem is that energy is transformed between two different forms. Because the rods have the potential to fall under the influence of gravity, they are said to have 'potential energy'. If the wheel turns, some of this converts to the 'kinetic energy' of movement. However, after one cycle, the rods return to their initial position, and therefore must have exactly the same potential energy (which is due to their position) as before. Since there is no external source of energy, and the rods have the same potential energy at every turn, there is nothing to put energy into turning the wheel.

Energy is conserved

By 1775, the Royal Academy of Sciences in Paris had had enough of perpetual motion. It issued a statement declaring that the Academy 'will no longer accept or deal with proposals concerning perpetual motion'. And in 1841, scientists finally found a scientific principle to throw at perpetual motion seekers: the first law of thermodynamics.

It was the first explicit statement of the conservation of energy. Leonardo da Vinci had suggested that, 'Falling water lifts the same amount of water, if we take the force of the impact into account,' but it took the German physicist Julius Robert Von Mayer

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 60

to explore the matter properly and issue an edict. Energy, he said, cannot be created or destroyed.

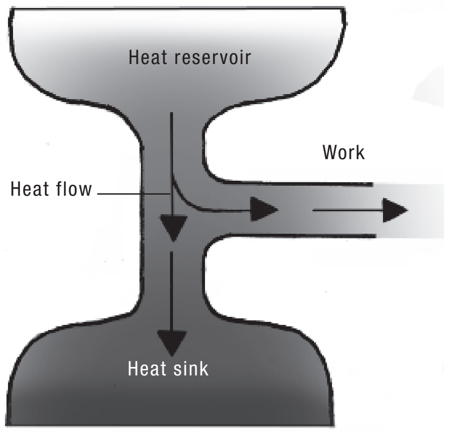
Not that he was taken seriously straight away: Von Mayer was told, for instance, to find some experimental evidence to back up this strange idea. This he did, by showing that the kinetic energy of vibration could be transferred to water molecules, manifesting as an increase in temperature. Once the point was proven, the principle was quickly accepted by physicists, and used to keep perpetual motion at bay. Motion takes energy, and the conservation of energy principle tells us that you can't get more energy out of a closed system than is there in the first place. Since friction affects any and every mechanism, dissipating some of that energy as heat and sound, inventing perpetual motion machines of the first kind became a fool's errand. Not that this put the perpetual motion seekers off. Around this time, the science of thermodynamics was giving them a whole new lease of life. Their goal? Perpetual motion machines of the second kind.

Miracle machines

The second kind of perpetual motion machine is something that extracts heat energy from a reservoir, such as the air or the ocean, and converts it into mechanical energy. It certainly seems like a good idea. The oceans are so vast a resource that, if we could extract heat that would cause a one degree drop in ocean temperatures, it would supply something like the energy needs of the United States for half a century.

The plausibility of this kind of machine is enticing. Indeed, creating an efficient steam-powered engine has been a human obsession since Hero of Alexandria created the 'aeolipile' in AD 1. This ball, that was set rotating by jets of steam, had no particular uses. However, subsequent inventions used steam turbines to turn spits, pump water from mines and power grinding pestles. None of them got anywhere near a truly useful efficiency, however. That efficiency came with James Watt's steam engine, first demonstrated in 1765. It was a development of the engine invented by Thomas Newcomen, and raised the efficiency

enough to kick off the Industrial Revolution. The theory behind such engines, though, was still very much in development. The builders of steam engines were working on hunch and intuition, not scientific theory.



THE CARNOT ENGINE

It wasn't until 1824 that the French military scientist Sadi Carnot published 'On the Motive Power of Fire'. Even then, this primary work in the field went largely unnoticed for a decade. But the scientific principles behind the steam engine were now in place. And, as a bonus, Carnot had worked out the principle that denies a free lunch to perpetual motion machines of the second kind.

There is a good reason why you can't get useful work out of a room temperature heat source. It is called the 'second law of thermodynamics', and it says, essentially, that you can't take the heat from something then turn all the heat into mechanical work. Some of that heat has to be passed on to a 'heat sink' at a lower temperature. It is the temperature difference between the heat source and the heat sink that determines how much work you'll get out of this 'heat engine'. Carnot showed that creating a perfectly efficient heat engine is impossible.

The rule of law zero

To see why, let's imagine an engine. Any engine seeking to perform work requires energy, which we will consider to come in the form of heat. Heat flows from a hot source to a colder one (this principle seems so obvious it was only formalized as the 'zeroth law' of thermodynamics long after the other laws were laid down), so both reservoirs are required; work can be extracted as heat flows from a hot 'reservoir' to a cold one.

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 62

The work extracted in this situation is the difference between the heat flowing out of the hot reservoir and the heat flowing into the cold reservoir. A perfect efficiency would have zero heat flowing into the cold reservoir so that all of the heat energy is used for the work you want to do.

Now let's consider, as Carnot did, the practicalities of the engine. Carnot imagined a piston engine much like the cylinder of a car engine, where the heat is used to expand gas that pushes on a piston. The gas is then compressed, and the cycle begins again. By considering the gas laws that relate pressure, temperature and volume, Carnot showed that the efficiency of an engine depends upon the ratio of the temperatures of the hot and cold reservoirs. No matter what fluid or gas is being used to power the engine, the ratio of the two temperatures is everything. And here is the problem with this free lunch.

The average diesel engine operates at around 550 Celsius. The exhaust gases exit to the outside temperature. The maximum efficiency possible, according to Carnot's work, is around 60 per cent. In reality, a diesel-powered car converts around 50 per cent of its fuel's chemical energy into energy that can move the car along the road. The rest is wasted as heat (which is why cars need cooling systems). Petrol engines are significantly less efficient.

What if we operate the two reservoirs at the extremes of temperature? In theory, the hot reservoir can operate at infinitely high temperatures. But the cold reservoir cannot be colder than absolute zero. Even dumping the heat in outer space would give a cold reservoir temperature of 3 K, or –270 degrees Celsius. Because you can't get lower than zero, and an infinitely hot reservoir does not exist (at least not one that we know about), a perfectly efficient engine is impossible. You cannot convert heat into work without wasting some of the heat. And that means that, to continue the cycle, you always have to put in energy. No free lunch, in other words.

WHY IS THERE NO SUCH THING AS A FREE LUNCH?

OUR UNIVERSE: THE ULTIMATE FREE LUNCH

According to physicist Alan Guth, there is such a thing as a free lunch. And we're living in it. The universe, Guth says, is 'the ultimate free lunch'. Guth is the originator of an idea in cosmology called 'inflation'. According to Guth, the universe, and all the energy it contains, seems to have arisen from little more than a gram of material. A fraction of a second after the big bang, the universe was a 100 billion times smaller than a proton, but it then blew up like a balloon. In fact, it blew up like a pea expanding to the size of the Milky Way in less time than it takes to blink an eye.

The numbers involved are staggering. It started when the universe was around one billionth the size of a proton. 10^{-34} seconds later it had expanded to 10^{25} times its original size – something around the size of a marble. And during this process, cosmologists reckon the energy within the universe increased by a factor of 10^{75}. It sounds like a violation of the something from nothing, or no free lunch, rule. But there's a complication that keeps it within the laws of physics: some of it is negative energy.

According to general relativity, our best description of the nature of space and time, the energy of a gravitational field is always negative. During inflation, the energy in the rapidly expanding space–time becomes ever more negative. Within this space–time, however, matter began to appear. That's because the properties of space–time mean that a portion of it spontaneously moves to a lower energy state: particles such as electrons, positrons and neutrons. Matter has positive energy, and the continuing creation of matter created more and more positive energy to balance the growing negative energy. The total energy can thus remain constant. The ancient Greeks said that nothing can be created from nothing, but inflation begs to differ.

Carnot's work led directly to the formulation of the second law of thermodynamics. As phrased by the English physicist Lord Kelvin and the German physicist Max Planck, it states that an engine operating in a cycle cannot transform heat into work without some other effect on its environment. Thanks to the second law, not only can you not get a free lunch, you can't even

001-065 Big Qs Physics .qxd:Big Questions 1/9/09 10:48 Page 64

> *'An engine operating in a cycle cannot transfer heat from a cold reservoir to a hot reservoir without some other effect on its environment.'*
>
> RUDOLPH CLAUSIUS

keep your lunch cool in the refrigerator for free. Refrigeration, it turns out, is nothing more complicated than the Carnot engine working in reverse.

In 1850, the German physicist Rudolph Clausius rephrased the Second Law to read, 'An engine operating in a cycle cannot transfer heat from a cold reservoir to a hot reservoir without some other effect on its environment.' A refrigerator, in other words, needs energy put into it. This arises from the natural tendency of energy to flow 'downhill': from hot to cold. Keeping the inside of your refrigerator below the temperature of your kitchen involves the same process of expanding and contracting, heating and cooling gases as running your car engine, and it all takes energy. This time, though, you need a compressor rather than an expander for the gas.

The march of entropy

As mentioned, Carnot's work involved consideration of the pressure, temperature and volume of the gas. The process that Carnot uncovered led to another revelation for physicists: the notion of entropy. The whole universe, it turns out, is spiralling into ever-more disorder. It was Clausius who classified this disorder as 'entropy', a word derived from the Greek for 'transformation'. In 1865 he wrote a mathematical treatise on the work that the atoms do on one another in a gas. The result, Clausius showed, is that the second law can be expressed in a new way: the entropy, or disorder, of a closed system either stays the same or increases – it never decreases.

That doesn't mean you'll never see entropy increase on a small scale. Your lunch inside the refrigerator will get cold, for instance, decreasing the disorder in its constituent molecules. But don't be fooled that this breaks the second law of thermodynamics. The inside of your fridge is not a closed system – the molecules of refrigerant gas take the heat away, and their

disorder increases as they do. As the heat is transferred to the air in your kitchen, the disorder in your house increases too.

This kind of thing is happening throughout the universe as the processes of nature unfold. It creates, in physicists' view, the irreversibility of natural processes: the arrow of time is just another way of expressing the second law of thermodynamics. The wasted energy of Carnot's engine cycle is the slow unravelling of the universe in microcosm.

Together, the first and second laws of thermodynamics put up a brick wall to any claims for the generation of a free lunch. So well proven are they, in fact, that the US Patent Office warns anyone submitting a patent for a perpetual motion machine that they should think carefully; they will most likely lose their money. 'The views of the Patent Office are in accord with those scientists who have investigated the subject and are to the effect that such devices are physical impossibilities,' the office's official statement says. 'The position of the Office can only be rebutted by a working model. . . . The Office hesitates to accept fees from applicants who believe they have discovered Perpetual Motion, and deems it only fair to give such applicants a word of warning that fees cannot be recovered after the case has been considered by the Examiner.' So not only is there no such thing as a free lunch. Even looking for one could end up costing you money.