A Kindle Single

DISCOVERING GRAVITATIONAL WAVES

JOHN GRIBBIN

Discovering Gravitational Waves

John Gribbin

Introduction

In February 2016, a team of researchers working at the Laser Interferometer Gravitational-wave Observatory (LIGO) in North America reported that they had made a direct detection of gravitational waves – ripples in space produced by the collision of two black holes far away across the Universe. This extraordinary experiment measures ripples much smaller than the size of an atom, and ranks as one of the most successful achievements in the whole of science. In one sense, the discovery was beautifully timed, coming exactly a hundred years after the publication of the scientific papers in which Albert Einstein, with his general theory of relativity, explained gravity in terms of distortions in spacetime. But in another sense, the announcement was ill-timed. Nominations for the Nobel Prizes close each year on 1 February, so the team were a couple of weeks too late to be considered for the 2016 Physics Prize. It is, though, impossible to think of a more worthy contender for the 2017 Prize. In this essay, I want to explain why the discovery is so important, and fill in the background to the story, starting with Einstein's masterpiece, completed at the end of 1915 and published in all its glory in 1916.

From Doubt to Certainty

Einstein did not mention gravitational waves in his great paper on the general theory but he had been thinking about them for several years because of the question of how fast a gravitational influence reaches out across the Universe. We can look at this in terms of the way the presence of any object that has mass distorts space around it (it distorts space and time, spacetime, but space is the bit we are interested in). The classic example (skip this if you know it) is to think of space as like a stretched rubber sheet, or the surface of a trampoline, and represent a mass like the Sun as a bowling ball plonked on the trampoline. The ball makes a dent on the trampoline, and marbles rolled across the surface follow curved lines (trajectories) around the dent. Similarly, the curved space around a massive object such as the Sun makes objects follow curved trajectories near it; we see this as a result of a force, which we call gravity. Even light is affected in this way, and it was Einstein's prediction of how much starlight would be bent as it passed near the Sun that enabled astronomers to test his ideas during a solar eclipse in 1919. The theory passed the test with flying colours, and Einstein became famous.

But what if the bowling ball is taken away? The curved surface of the trampoline becomes flat again, but it doesn't do so instantly. The smoothing out spreads across the surface. The Earth is following an orbit around the Sun because of the dent the Sun makes in spacetime. I have to be careful here not to say that it is just a dent in space; the situation is a bit more complicated than the bowling ball analogy, but I shall spare you the details. If the Sun suddenly ceased to exist, would the Earth instantly fly free in space, or would the curve still be there for a while, at the distance of the Earth, holding the planet in its orbit until spacetime flattened itself out, and only letting it fly free when there had been time for the information that the Sun had gone to reach us? The best guess was that gravity must 'propagate' in some way at the same speed as light. Light from the Sun takes just over 8 minutes to reach the Earth, so if the Sun were to go out, our sky would only go dark 8-and-a-bit minutes later. Since Einstein already knew, from the special theory he had published in 1905, that nothing could travel faster than light, he expected that gravity would travel at the same speed. So if the Sun disappeared, the Earth would continue in its orbit and the sky would be bright for another 8-and-a-bit minutes, then the sky would go dark and the planet would fly free at the same time.

But think again about that bowling ball being taken off a trampoline. The stretched surface doesn't instantly go back to being flat; it bounces up and down for a while as it settles down, sending ripples across the surface. If the Sun disappeared, the space (spacetime) around it would presumably ripple in the same way, with the ripples dying down while it smoothed out. These ripples would be gravitational waves.

Later in 1916, Einstein did publish a paper on gravitational radiation (in this context, the terms 'waves' and 'radiation' are interchangeable, but not 'gravitational' and 'gravity; gravity waves are the waves made by a fluid, such as your bathwater, sloshing about under the influence of gravity). But even Einstein was fallible – in fact, he was not very good at maths, by the standards of Nobel-prizewinning physicists – and he made a mathematical mistake in that paper which led to a false conclusion which confused the issue. It took two years for his colleagues to convince him that he had indeed made a mistake, and for him to accept the correct description of the waves described by the general theory of relativity. So 1918 was really the year in which gravitational wave theory got off the ground. But unlike the light-bending prediction, there was no obvious way to test whether gravitational waves really existed. You can't make the Sun disappear to test the theory. Einstein himself was never quite sure if it was right. According to the theoretical cosmologist Janna Levin he once said, 'If you ask me whether there are gravitational waves or not, I must answer that I do not know. But it is a highly interesting problem.' It would be nearly sixty years before all doubts about the reality of gravitational waves were removed by the discovery of an astronomical object known as 'the binary pulsar'.

After an initial burst of attention for Einstein's theory following the success of the light-bending observations, the general theory became overshadowed by the success of the quantum theory developed in the 1920s and beyond. Quantum physics had practical implications in things such as chemistry, explaining how molecules work, and electronics, leading to the development of transistors and other solid state devices. But the general theory didn't seem to have any practical applications, and was largely left to mathematicians who enjoyed tinkering with the equations. The one area of science where it did apply was in cosmology, the investigation of the Universe at large, where it turned out that the equations of the general theory describe an expanding Universe exactly like the one revealed by observations of distant galaxies. But right up until the 1960s, and the discovery of the cosmic background radiation that is interpreted as the 'echo of the Big Bang', only a few people took these cosmological calculations seriously, and the ones who did argued about what they meant; there was even some uncertainty about just how general the general theory really was. . As late as 1966, when I told the cosmologist Herman Bondi that I wanted to study cosmology, he advised me that if I wanted to be taken seriously as a scientist I should study something 'more respectable' first. Mathematicians loved it, but physicists, for the most part, just didn't bother with it. As for gravitational waves, like Einstein they thought them to represent an interesting problem to which nobody knew the answer.

The exception to the scepticism (if that is not too strong a word) about the general theory was a small band of scientists, astrophysicists who were interested in what happens to stars when they run out of fuel. In the prime of its life, a star keeps shining because deep in its heart the nuclei of lighter elements are being fused to make heavier elements – hydrogen to make helium, helium to make carbon, and so on. At each stage, a little of the mass (m) involved is released as energy (E) in line with Einstein's famous equation (from the special theory), $E = mc^2$. The release of this energy produces an outward pressure which holds the star up against the inward tug of gravity. But the process stops when the 'fuel' has been converted into stable nickel and iron, and even sooner for less massive stars (including our Sun). The details (discussed in my book *Stardust*) are not important here; the thing that interested astrophysicists was what happens when no more energy is released in its heart, and a star has to collapse down upon itself as the pressure inside is reduced.

What happens depends on nuclear processes (essentially, quantum theory) and gravitation (essentially, the general theory). Nuclei are crushed together by gravity, but resist as they are pushed together more tightly, like marbles being squashed in a bag. Atomic nuclei are made of positively charged protons and electrically neutral neutrons. In an atom, a single nucleus is surrounded by its own set of negatively charged electrons, which balance the charge of the protons to make each atom electrically neutral. But inside a star the nuclei swim about freely in a sea of electrons. A cooling star with no internal source of energy, known as a white dwarf, will shrink down until all the nuclei are in close proximity, but still in a sea of electrons, being held up against further collapse by a resistance known as quantum degeneracy.

To put this in perspective, when our Sun, which is now a bit more than a hundred times bigger across than the Earth, becomes a white dwarf it will be about as big as the Earth is today. But for many stars this is not the end of the story. In the 1930s the astrophysicist Subrahmanyan Chandrasekhar showed that if a stellar remnant has more than about 1.4 times the mass of our Sun (which is now known as the Chandrasekhar limit), gravity overwhelms the quantum forces, and the star will shrink even further. In essence, the protons and electrons are squeezed together to make neutrons, so all that is left is a ball of neutrons, like a single giant atomic nucleus. Such a neutron star might contain a couple of times as much matter as our Sun, but would be no bigger across than a few kilometres – about the size of a large mountain on Earth. But there was considerable scepticism about whether such neutron stars could really exist, and some physicists thought that there must be an undiscovered law of nature which would prevent gravity pushing things to such extremes. That all changed in 1967.

That year, a graduate student called Jocelyn Bell (now Jocelyn Bell Burnell), working at the University of Cambridge, discovered a previously unknown kind of astronomical object, stars which produce rapid 'pulses' of radio noise with a very precise beat. They soon became known as pulsars, and the nature of the pulses showed that they had to be smaller than the Sun. The first guess of astrophysicists was that these might be white dwarf stars that were literally pulsating, 'breathing' in and out in a regular way. But computer simulations of the way white dwarf stars oscillate soon showed that a white dwarf could not oscillate rapidly enough to explain the observations without blowing apart. These calculations (in which I was involved in a small way) left only one possibility – pulsars must be neutron stars. Specifically, they must be highly magnetic neutron stars from which beams of radio noise are being swept around the sky like the beams of light from a lighthouse, flicking past the Earth regularly and precisely as the star spins. This discovery showed that gravity could indeed overwhelm quantum degeneracy, and soon produced a revival in the study of the general theory and the behaviour of gravity under extreme conditions, more of which shortly. But the study of pulsars also produced the first direct evidence that gravitational waves really do exist.

The observations that proved gravitational waves are real were made using a giant radio telescope – which featured in the movie *Contact*, starring alongside Jodie Foster – at Arecibo, in Puerto Rico. This dish, a thousand feet (305 metres) across, was for many years the largest radio telescope in the world, until the completion of an even larger dish in Guizhou Province in south-west China in 2016. It is impossible to build a fully steerable radio telescope, like the famous Jodrell Bank dish, as big as the Arecibo reflector, because it would collapse under its own weight. So it is built into a natural depression 50 metres deep and covering more than 7 hectares in the hills of Puerto Rico. This makes it possible to support the weight of the dish from underneath. Although this means that the dish itself cannot be moved, the receiving apparatus is suspended above the dish on wires from three tall pillars, like the supports of a suspension bridge, and can be shifted slightly to look at different parts of the sky. So in effect the beam can be swung 20 degrees either side of the vertical, enabling the telescope to study radio emissions from anything in a broad band of the sky as it is carried round by the rotation of the Earth. This is no mean feat of engineering – the receiver weighs 900 tonnes and is suspended 150 metres above the dish, where it has to be moved sideways by as much as 50 metres, but to an accuracy of about a millimetre.

The Arecibo dish was initially completed in 1963 (it has since received several upgrades) and made the dramatic observations confirming the existence of gravitational waves eleven years later. The telescope is almost perfect for studying pulsars, provided they are in the right part of the sky, and in 1974 Russell Hulse, a Ph.D. student at Harvard University, was using the Arecibo telescope to carry out a search for pulsars, under the supervision of Joseph Taylor. As the junior member of the team, Hulse was the one who had to be on site in Puerto Rico, while Taylor was mostly concerned with his academic duties back home. This was one of the first such surveys to be computerised, using a then state-of-the-art minicomputer to sift the radio noise coming in to the telescope and identify pulsar candidates, which the astronomers could then study in more detail to reject spurious signals. In a fourteen-month stay at Arecibo, from December 1973 to January 1975, Hulse and the minicomputer identified forty previously unknown pulsars. This was a major achievement in its own right; but one of those pulsars turned out to be something special.

On 2 July 1974 the computer indicated that it had found a pulsar right at the limit of its capacity to make identifications. If it had been any fainter, it would not have been spotted. After observing the faint source for several weeks, Hulse confirmed that it was not a false alarm, and dubbed the object PSR 1913+16. The 'PSR' is shorthand for pulsar, and the numbers indicate the position of the object in the sky, like latitude and longitude on Earth. It was a particularly exciting discovery because the beam from the pulsar was sweeping round so fast that it produced 17 blips every second, meaning that the neutron star was spinning once every 58.98 milliseconds, making it the second-fastest pulsar known at the time. In those days before the advent of email, Hulse dashed off an airmail letter to his supervisor to tell him the good news. Then, towards the end of August, he got down to studying the 'new' pulsar in more detail.

What he found had him scratching his head. In order to pin down the exact period of the pulsar, he made measurements one hour apart, intending to combine them into a single accurate measurement. But he found that in an hour, the period of this particular pulsar had changed by 27 milliseconds, which was an unheard of ‘error’ for a pulsar. At first, Hulse thought that he was going wrong somewhere. But measurement after measurement showed that the pulsar’s period really was changing. Sometimes it increased by a few milliseconds, sometimes it decreased. Hardly able to believe what he had found, Hulse realised that the changes were caused by the pulsar orbiting around another star – or rather, that it is a member of a binary pair, with each star orbiting around their common centre of mass. When the pulsar is moving towards the Earth, the pulses that give it its name are piled up closer to one another; when it is moving away, the gap between them is stretched out. This is a variation on the well-known Doppler effect. And the speed with which the changes are occurring showed that the object that PSR 1913+16 is orbiting round must be another neutron star. This isn’t visible to us as a pulsar, because its ‘lighthouse beams’, if it has any, are not directed our way. When Taylor received this additional news, he dropped everything, handing over his teaching duties to a colleague, and hurried to Arecibo to work with Hulse in determining the details of this two-star system; it became known as ‘the binary pulsar’, although only one of the neutron stars is visible as a pulsar.

Together, Hulse and Taylor determined the ‘orbital parameters’ of the system. The pulsar zips round its companion once every 7 hours 45 minutes, reaching a maximum speed of 300 kilometres per second (a thousandth of the speed of light), with an average speed of 200 kilometres per second. The size of the orbit traced out at this speed is about 6 million kilometres, which just happens to be about the same as the circumference of the Sun. So if the orbit were circular the whole binary pulsar system would fit inside the Sun, with the two neutron stars separated by a distance equivalent to the radius of the Sun. In fact, the orbits are elliptical, and the minimum separation of the two objects (periastron) is about 1.1 solar radii; the maximum separation (apastron) is 4.8 solar radii. The astronomers immediately realised that such a system could act as a test bed for the general theory, and in particular that it ought to be producing gravitational waves, which would have an effect on the orbit. But it would take years to measure such an effect, and Hulse, who had more than enough material to earn his Ph.D., went on to other work, while Taylor and other colleagues made regular visits to Arecibo to monitor the changes in the orbital parameters – in particular, the orbital period – year by year.

One of the first things the team were able to measure was the way the orbit of the pulsar seems to shift sideways with every circuit, so that it doesn’t trace out a simple ellipse but a pattern like a child’s drawing of a daisy. Even without being able to see the binary pulsar, this shift can be inferred from the way the pulsar was seen to speed up and slow down in its orbit. The same thing happens with the planet Mercury, and is predicted by the general theory, but not by Newton’s theory of gravity. Explaining this ‘advance of the perihelion’ of Mercury was one of the first triumphs of Einstein’s theory. The equivalent for the pulsar is called the advance of the periastron, and provided an even more accurate test of the general theory, because Mercury only orbits the Sun four times a year, but the binary pulsar orbits its companion a thousand times a year. Combined with the other observations, the shift in the periastron showed that the combined mass of the two stars in the binary is 2.8275 times the mass of our Sun. But this was just the beginning. Data piled up with impressive accuracy. At the beginning of September 1974, the period of the pulsar itself was 0.059029995271 seconds, increasing at a rate of 0.253 nanoseconds per year as the spin of the star slowed down; the orbital period was 27906.98163 seconds; and the rate of change of the periastron was 4.2263 degrees of arc per year. The timekeeping of the pulsar was affected by influences described by both the special and the general theories of relativity. Because it moves at high speed, the ‘clock’ of the pulsar is affected in one of the effects described by the special theory, which tells us that moving clocks run slow. Because it is moving in a strong gravitational field, the general theory tells us that the clock will be affected by an additional process, because clocks in a strong gravitational field run slow. Both effects vary over an orbit, as the neutron star moves faster or slower, and in and out of the region of strongest gravity. (Incidentally, all this is of more than academic interest. Both these effects operate at a more subtle level to affect ‘clocks’ on satellites in orbit around the Earth. They have to be taken account of in the workings of the GPS network which gives the information that enables your smartphone to tell you exactly where you are, and to give you a route to somewhere else. Millions of people use the general theory of relativity every day without knowing it!)

Putting everything together, the measurements of the way the ‘ticking’ of the pulsar changes over an orbit could be used to determine the ratio of the masses of the two stars, and plugging this back into the measurement of the total mass of the system revealed that the pulsar has a mass 1.44 times that of the Sun, while its companion has a mass just under 1.39 times the mass of the Sun. All this determined for a system some 16,000 light years away; but it was just the background to the measurement of the changes caused by the emission of gravitational radiation from the system.

Imagine two hollow metal spheres, connected by a short rod, floating in a tank of water. If they do not move, there will be no waves in the water. But if they are rotating around one another, like a spinning dumbbell, waves will ripple out across the surface. According to the general theory, that is similar to what happens to spacetime if two neutron stars separated by less than the diameter of the Sun are orbiting around one another. And with all the accurate numbers provided by the observations, the general theory could be used to predict how much energy would be going in to the ripples. If energy from the binary system is going into the gravitational ripples, the two stars have to slowly spiral together to compensate, moving faster as they do so, so the orbital period will decrease by a tiny amount which can be precisely calculated.

The prediction for the binary pulsar was that the orbital period, which is about 27,000 seconds, would decrease by about 0.0000003 per cent, or 75 millionths of a second, each year. In order to measure such a tiny effect, the astronomers had to make allowances for all kinds of disturbances, such as the motion of the Earth in its orbit around the Sun and tiny changes in the rotation of the Earth itself – the kind of changes that lead to the introduction of 'leap seconds' into our calendars from time to time. After taking all these influences into account and analysing some five million pulses from PSR 1913+16, in December 1978 Taylor was able to announce the result to an international meeting in Germany. The orbit of the binary pulsar was 'decaying' exactly in line with the predictions of the general theory. This was spectacular confirmation that the general theory is right, and gravitational waves are real. In 1993, Hulse and Taylor shared the Nobel Prize in physics for their discovery.

As the binary pulsar has continued to be studied, details of its behaviour have been refined. A few of the numbers are worth mentioning. The total power radiated by the system in the form of gravitational waves is 7.35×10^{24} watts, just under 2 per cent of the power radiated by our Sun in the form of light. The Solar System, by comparison, is calculated to be radiating about 5,000 watts in the form of gravitational waves, mostly produced by the effect of Jupiter orbiting the Sun. The rate of decrease of the orbital period of the binary pulsar is 76.5 microseconds per year, the size of the orbit (semimajor axis) is shrinking at 3.5 metres per year, and the two neutron stars will collide in about 300 million years.

Once they knew what to look for, astronomers soon discovered more objects like PSR 1913+16. More than fifty binary pulsars are now known, although the one discovered by Hulse and Taylor is still referred to as ‘the’ binary pulsar. Most of them are binaries in which a pulsar is orbiting a white dwarf star, making them slightly different from the Hulse–Taylor binary, so I will mention one example of this variety, which is also an excellent test bed for the general theory. PSR J0348+0432 is a system 7,000 light years away in which a pulsar is the companion to a white dwarf. The two stars orbit around each other once every 2.46 hours, while the pulsar spins on its axis roughly 25 times a second (once every 39 milliseconds). The now standard analysis shows that in this case the neutron star has a mass just over twice that of the Sun, while the white dwarf has a mass less than a fifth of the mass of the Sun, comfortably below the Chandrasekhar limit. The average distance between the two stars is about 1.2 times the radius of the Sun, so the whole system would fit inside the Sun, and the orbit is calculated to be decaying as a result of gravitational radiation at a rate of 2.6×10^{-13} seconds per second. The measured rate is 2.7×10^{-13} seconds per second, or eight millionths of a second per year, agreeing with the prediction of the general theory within the (small) uncertainty of the measurements. The decay is, of course, actually measured over several years, from which the ‘per second’ rate is calculated.

By the end of the 1970s, there was no room to doubt that gravitational waves exist. But that left an enormous challenge, which few people thought could be overcome – to detect gravitational waves, ripples in space, directly, here on Earth. If this could be done, it would open up a new window on the Universe, a way to study what is going on ‘out there’ without using light, radio waves, or any other kind of electromagnetic radiation. A handful of people had taken up this challenge even before the discovery of the binary pulsar. But it was the evidence that gravitational waves are real that took such projects to a higher level and, not least, helped to ensure funding for them. But what should they be looking for? What kind of astronomical event could produce ripples in space that would be detectable on Earth, thousands (or millions) of light years away? It had to be something involving objects even more extreme than neutron stars – black holes.

The Hole Truth

Black holes were another prediction of the general theory that for decades nobody quite knew whether or not to take seriously. This time the prediction was not made by Einstein, but by another German, the astronomer Karl Schwarzschild. Schwarzschild, who was serving in the German army at the time, heard about the general theory from Einstein himself, and wrote two scientific papers about its implications which Einstein presented to the Prussian Academy of Sciences for him in January and February 1916 (actually before Einstein published his work on gravitational waves). Schwarzschild died of a rare skin disease in May that year, but those two papers ensured his place in the scientific pantheon. What he had found was that if any amount of matter is squeezed into a small enough volume it will bend spacetime around itself so much that it will be cut off from the outside world. Things could still fall on to (or in to) the cut-off region of space, but nothing, not even light, would be able to escape. The boundary between the cut-off mass and the Universe outside would, in the simplest case, form a spherical surface at a certain radius, which became known as the Schwarzschild surface and the Schwarzschild radius. The Schwarzschild radius is bigger for bigger masses, but for the kind of things astronomers knew about in 1916, stars and planets, it is very small. The Schwarzschild radius for an object with the mass of the Sun is 2.9 km, and the Schwarzschild radius for an object with the mass of the Earth is 0.88 cm. Nobody could imagine a process which could squeeze a star into a ball less than 3 km across, or anything that could squeeze a planet into a ball less than a centimetre across. So these hypothetical objects were regarded as a mathematical curiosity, something which the equations said might exist under extreme conditions, but which nobody believed really did exist in the Universe in which we live.

Things started to change in the 1930s, when Chandrasekhar discovered the limiting mass for a white dwarf star, but even then the implications were slow to sink in. Chandrasekhar’s calculations were published in 1931, and the

Russian physicist Lev Landau independently discovered the limiting mass for a white dwarf at around the same time, without knowing about Chandrasekhar's work. Their calculations suggested that a stellar remnant with more than 1.4 times the mass of the Sun would collapse indefinitely, into a point of infinite density; but the conclusion drawn from this by Landau sums up the attitude of physicists at the time: 'We must conclude that all stars heavier than 1.5 solar masses certainly possess regions in which the laws of quantum mechanics . . . are violated.' Something, physicists felt, must stop the collapse, even if they did not know what it was. In 1932, though, the discovery of the neutron suggested to a few people that stable stellar objects even more dense than white dwarf stars – neutron stars – might exist.

Just two years after the discovery of the neutron, in 1934 Walter Baade and Fritz Zwicky, working in the USA, pointed out that if a dying star collapsed into a ball of neutrons an enormous amount of gravitational energy would be released, sufficient to explain the stellar outbursts known as supernovas. In a passage drafted by Zwicky (Baade had doubts), they commented:

> With all reserve we advance the view that a super-nova represents the transition of an ordinary star into a *neutron star*, consisting mainly of neutrons. Such a star may possess a very small radius and an extremely high density . . . A neutron star would therefore represent the most stable configuration of matter as such.

This idea was way ahead of its time, and to almost everyone in the 1930s, literally unbelievable, although it is now known to be very close to the truth. In terms of radius, a white dwarf star is about one-hundredth the size of the Sun, but a neutron star is one seven-hundredth (0.0014 times) the size of white dwarf – 0.000014 times the size of the Sun. From another perspective, a white dwarf is a couple of thousand times bigger than the Schwarzschild radius for its mass, but a neutron star 10 km across would be only three times bigger than the corresponding Schwarzschild radius – too close for comfort. Astronomers shied away from contemplating such objects, but one team of American physicists was sufficiently intrigued to try to find out if, assuming neutron stars really might exist, there was a limiting mass for neutron stars, analogous to the Chandrasekhar limit for white dwarf stars.

This involved calculating the behaviour of a ball of neutrons (its so-called equation of state) under extreme conditions. It is no coincidence that the person who was intrigued enough to try working out the equation of state of matter under such extreme conditions was Robert Oppenheimer, remembered now as the 'father of the atom bomb'. The physics involved in exploding nuclear bombs is very similar to the physics involved in working out the equation of state for a neutron star. In a paper published with his student George Volkoff in 1939, Oppenheimer suggested that stable neutron stars could exist only if they had masses in the range from 10 per cent to 70 per cent of the mass of the Sun; any less mass and their self-gravity would not be strong enough to hold them together, but any more mass and gravity would overwhelm quantum effects and they would collapse to a point. As a better understanding of the equation of state was developed, these numbers were revised to set an absolute upper limit on the mass for such a star of three solar masses, now known as the Oppenheimer–Volkoff limit. For greater masses, there is no way to hold the star up against the inward tug of gravity. But Oppenheimer and Volkoff offered a way out of the dilemma, suggesting that the gravitational distortion of spacetime around such an extreme object would make time run so slowly that the collapse would almost come to a halt, that 'the star will continue to contract indefinitely', but 'one would hope', they wrote, that 'the rate of contraction, and in general the time variation, [would] become slower and slower'.

It was, though, the investigation of collapsed stars that came to a halt around this time, as people like Oppenheimer were diverted into war work. After the war, nobody picked up the baton, and nobody except Zwicky really believed that neutron stars existed anyway. But everything was shaken up in 1967 with the discovery of pulsars and their identification as rapidly spinning neutron stars. A few mathematical physicists who had been interested in using the equations of the general theory to describe extreme conditions, whether or not such extreme conditions existed in the real Universe, suddenly found themselves at the centre of astrophysical attention, and were in demand to explain the behaviour of actual stars. One of those relativists, John Wheeler, building from Oppenheimer's work, was already interested in the equations that described what were then cumbersomely known as 'gravitationally completely collapsed objects'. When pulsars were discovered, a meeting was held at NASA's Goddard Institute for Space Studies to consider all possible (and some impossible) candidates for them. Everything was up for discussion, including gravitationally completely collapsed objects and the possibility of signals from aliens. According to Wheeler, he was giving a talk suggesting that pulsars might be 'gravitationally collapsed objects', when somebody in the audience called out 'Why don't you call them black holes?' The name stuck. Although pulsars turned out not to be black holes, the discovery of objects only three times bigger than black holes helped to drive a wave of investigations by theorists into how black holes might behave. And just a few years later they had observational evidence for the existence of at least one black hole, only about 6,000 light years away from us. The discovery came from the opening-up of another new window on the Universe, worth describing in a little detail because it shows the kind of thing we might expect from the opening of the gravitational wave window.

In the 1960s, the discovery of strong X-ray sources in space came as a complete surprise.. The Sun was known to be a weak emitter of X-rays, detected by instruments carried into space by early rockets. Some theorists suggested that energetic particles from the Sun (cosmic rays) might strike the surface of the Moon and energise atoms there to make them emit X-rays. An attempt to test this suggestion was made in June 1962, using a rocket to lift detectors above the atmosphere of the Earth, which absorbs X-rays, for a mere six minutes. The experiment did not see any X-rays coming from the Moon. But as the rocket rotated and the field of view of the detectors scanned around the sky, they picked up a bright source of X-rays coming from a point on the sky. Later rocket flights confirmed the existence of this source, always in the same place, in the direction of the constellation Scorpius. It was dubbed Sco X-1, and later work (which I was involved with in a minor role) showed that the X-rays are emanating from a neutron star on to which matter is falling from a companion star, releasing gravitational energy which powers the production of X-rays.

Other rocket flights in the 1960s revealed a few other X-ray sources, but X-ray astronomy really took off in 1970 when a satellite dedicated to X-ray observations was launched. Instead of being limited to observing the skies for a few minutes, the detectors on board this satellite, which was known as Uhuru, could keep scanning the skies as long as the satellite was in orbit and its power lasted – which turned out to be three years. Uhuru found that the sky is covered with X-ray sources, some of which, like Sco X-1, could be identified with a visible star system, while others were seen only in X-rays. X-ray astronomy became a scientific discipline in its own right, but just one of the sources studied by Uhuru was enough to prove that black holes exist.

Sco X-1 is the brightest X-ray source on the sky, as seen from Earth. The second brightest lies in the direction of the constellation Cygnus, and is known as Cyg X-1. It had actually been seen on the same rocket flight that discovered Sco X-1, and like Sco X-1 it ‘flickers’ rapidly in X-rays, brightening and dimming repeatedly. The speed of this flickering is related to the strength of the gravitational field in which the material emitting the X-rays is gripped. Faster flickering means stronger gravity. The flickering of Sco X-1 helped to prove it must be a neutron star; the flickering of Cyg X-1 was even faster, which meant that the X-rays were coming from something with a stronger gravitational pull than a neutron star. Could it be a black hole?

The problem was that there was no known star that could be identified with Cyg X-1. The X-ray observations were not precise enough to pin down the exact spot on the sky the radiation was coming from. All they could do was indicate a patch on the sky where the X-rays originated, and none of the stars in that patch of sky looked particularly interesting (by contrast, the optical counterpart of Sco X-1 is an unusual and interesting star). But radio astronomers found that the same patch of sky contained a radio source whose behaviour seemed to be linked with that of the X-ray source. Long-term monitoring of Cyg X-1 by Uhuru was compared with the observations of radio waves from the same part of the sky, and the combined set of data showed that the X-rays were coming from the vicinity of an ordinary-looking star that had been catalogued many years before by Harvard College Observatory and given the prosaic catalogue number HDE 226868. Little did they know what was being labelled in this way.

Immediately Cyg X-1 was identified with HDE 226868, optical astronomers around the world turned their telescopes on this star to analyse its behaviour. They determined that it is a member of a class known as blue supergiants (because, you guessed, they are big and blue), and that it is in a binary system, being tugged from side to side by an unseen companion. The blue supergiant is orbiting around the unseen companion once every 5.6 days, at a distance about a fifth of the distance from the Earth to the Sun. A blue supergiant cannot possibly have a mass less than 12 times that of the Sun, and most are about twice that mass. From the same basic laws of orbital mechanics used to work out the masses of binary pulsars, it was simple to calculate that with an orbital period of 5.6 days and a mass for the giant star of 12 solar masses, the companion must have three times the mass of the Sun. If the blue star is more massive, the companion must also be more massive. If the companion were another bright star held up by nuclear reactions going on in its heart we would see it. But it is not, and its mass is above the Oppenheimer–Volkoff limit. It could only be a black hole. The X-rays were then explained as emission from hot gas piling into the black hole under the influence of gravity. Further observations showed that the minimum mass of the blue star is 16 times that of the Sun, implying a black hole mass of 7 times that of the Sun, while the most likely mass for the blue star is 33 solar masses, implying a mass for the black hole of 20 times that of the Sun.

Many more such black holes have now been identified, and their existence has been explained in terms of the same supernova explosions that Zwicky invoked to predict the existence of neutron stars. Zwicky was an ebullient, larger-than-life character who was never afraid to make extreme projections from existing knowledge. But even he did not guess that some of these explosions are so violent that while the outer layers of the star are blasted away into space, the inner core is squeezed so hard that it becomes a black hole. This is the kind of event which ought to produce a burst of gravitational waves that might be detected on Earth, if the supernova occurred close enough to us. The objects produced in explosions like this are known as ‘stellar mass’ black holes, to distinguish them from the much larger ‘supermassive’ black holes that are now known to lie at the hearts of galaxies, including our own Milky Way galaxy. Those objects may contain millions, or even billions, of times the mass of our Sun. Their story is outside the scope of this essay, except that they are the kind of objects which might, before too long, be studied using gravitational waves. But just to put them in perspective, the relatively small supermassive black hole at the heart of the Milky Way has a mass of around four million times the mass of the Sun. We know this because we can follow the orbits of stars moving around it and work out how strongly they are being tugged by whatever is holding them in those orbits. The corresponding Schwarzschild radius is 12 million km, 17 times bigger than the radius of the Sun. The gravitational radiation detected by LIGO, however, involves objects much more like the black hole in Cygnus X-1, with masses a few dozen times the mass of the Sun.

The study of black holes is a saga in itself. What actually goes on inside a black hole? Could they be the entrances to tunnels through space and time? Nobody knows, although there has been a great deal of mathematical speculation. But that doesn't matter for now. What matters is that black holes do exist, and that they can interact with one another and the Universe at large through gravity.

Raising the Bar

The first person to attempt to detect gravitational waves here on Earth was an American physicist, Joseph Weber, working at the University of Maryland. Weber's background was in electronic engineering, but in the 1950s he became intrigued by the general theory of relativity and its predictions, in particular the possible existence of gravitational waves. He decided to build a detector to search for this radiation, starting out from the standpoint of an engineer. His philosophy, as he later told science writer Marcia Bartusiak, was to 'build something, make it work, and see if you find anything'. (Bartusiak's book is the best source of information about Weber's work, based on an interview with him. I am grateful to her for letting me re-use some of her material here.) It was an ambitious idea, because the predicted effects of gravitational waves here on Earth are so small. As a wave passes by, it first squeezes and then stretches space, which makes objects squeeze and stretch accordingly. As Bartusiak has spelled out, an event, such as the collision of two black holes, which produces an intense burst of radiation, might briefly stretch a nearby object to twice its original length; a human being two metres tall would be stretched to four metres, with unpleasant results. But if the same event occurred as far away from us as the centre of the Milky Way, an object one metre long here on Earth would be stretched by a billionth of a billionth (10^{-18}) of a metre, very roughly a hundred-millionth of the size of an atom. In engineering terminology, the strain would be 10^{-18} metres per metre. Such a strain would change the distance between the Earth and the Sun by the size of a bacterium. Yet Weber published a paper in 1960 setting out how it might be possible to measure such effects, then set out to build an experiment to do just that.

The key to Weber's idea was that although a gravitational wave passing through a solid metal cylinder would stretch and squeeze the cylinder only by a tiny amount, if the cylinder were just the right size this would set it ringing, briefly, like a bell struck with a hammer. This ringing would be more pronounced if the length of the bar matched the wavelength of the gravitational radiation, so in effect it would be tuned to respond to certain waves. This is like the way a string on a guitar leaning against a wall will resonate when a note that matches the fundamental note of the string is played on another instrument nearby. Weber's electronic expertise came in to play with the design of the instruments to measure very small vibrations of the bar – a ring of detectors around the waist of the cylinder to convert the vibrations into electric signals that could be recorded and analysed.

Weber's first detector was constructed with the assistance of Robert Forward, a graduate student who went on to become a top physicist, and also a leading science fiction writer. (I particularly recommend his book *Dragon's Egg*.)

It was a cylinder made of aluminium, five feet long. The size was not based on any cunning calculations of the behaviour of gravitational waves. Forward recalls that it was based on how big a cylinder he could grasp with his outstretched hands. Purely by chance, however, when neutron stars and black holes were discovered a few years later it turned out that the gravitational waves produced by them would, in theory, have wavelengths that ought to produce resonance in cylinders a few feet long. But that realisation still lay in the future when Weber's first bar detector was completed in 1962.

In those early days, Weber built several similar bars, all in the same location, but realised that they were being affected by local disturbances such as trucks rolling past. So he set up two identical bars, each of them five feet long and weighing 3,100 pounds, one installed at Maryland and the other 700 miles away at a lab near Chicago. He reasoned that only gravitational waves would affect both bars simultaneously (actually, with a tiny delay, assuming the waves travel at the speed of light), and they were linked by a telephone line to record any 'coincidences' – disturbances that affected both detectors within a 0.44-second time gap, or 'window'. Starting in December 1968, Weber and his team recorded several events that they could not explain away by any outside influences, in just the range that theory predicted waves would be produced by a star exploding, with its core settling down as a neutron star. This was just at the time interest in the general theory was being revived by the discovery of pulsars, and Wheeler was publicising the term 'black hole'. When Weber announced his results at a meeting of relativists in June 1969, they caused a sensation. Many other physicists were motivated to take up the search for gravitational waves, and several of them visited Weber's lab to study his equipment for themselves. With hindsight, one of the most significant of those visitors was Ron Drever, from the University of Glasgow.

Efforts to follow up Weber's apparent discovery were soon under way in the Soviet Union, Italy, Germany, Japan, and both England and Scotland. In the United States alone, gravitational wave experiments were developed at Bell Laboratories, the IBM research centre in New York, at the University of Rochester, and at Stanford University. As Tony Tyson, of Bell Labs, said, in a scientific paper published in 1972, 'It is clear that if it were not for Weber's work, we would not be as near as we are today to the possible detection of gravitational radiation.' A variety of detectors was devised with the dual aims of improved sensitivity compared with Weber's bars, and the ability to detect different wavelengths of radiation. But nobody found anything, even though Weber continued to report detections of what he thought were gravitational wave events.

The theorists had a field day, though. If the pulses being reported by Weber were coming from neutron stars, then their energy meant that the neutron stars had to be within about 300 light years of Earth, almost in our astronomical back yard. But the Maryland team was reporting roughly one event per day, and that required far more neutron stars than could possibly be that close to us. On the other hand, if the events were coming from the heart of the Milky Way, at the galactic core, they would have to be enormously energetic to be detected on Earth. Einstein's equation told the theorists how much mass would have to be converted into energy each day to explain the observations. It represented turning the equivalent of the mass of the Sun into pure gravitational energy every day. Losing mass at that rate would weaken the gravitational bonds that hold the Galaxy together; it would have dissolved away long ago, as gravity loosened its grip on all the stars. As for supernova explosions, to explain the observations there would have to be a thousand times more supernovas than optical observations revealed or theory predicted. The clinching evidence came from two European groups, one at Frascati, in Italy, the other at Munich, in Germany. They built detectors that were nearly identical to Weber's original design, and they were working up to the middle of 1974. They also found nothing.

At a meeting held in the summer of 1974, Drever, who had been running his own experiment with negative results, summed up the situation: 'I think that when you put all these different experiments together, because they are different, most loopholes are closed.’ But what Weber had started could not be stopped, especially with the discovery of neutron stars and black holes confirming the accuracy of the general theory as a description of what happens under extreme conditions. The gravitational wave genie was out of the bottle. Already, people were thinking of new, better and more sensitive ways to search for the waves, which by now everyone was sure existed. ‘Another technique,’ said Drever, ‘is the quite different possibility of having separate masses which are a long distance apart . . . One may monitor the separations using laser techniques.’ He was already looking to the future when ‘the thing could rapidly spread to where we would have a real astronomy and we would be producing maps of the sky of gravitational wave sources’. And as another contributor to the conference, Tony Tyson noted, ‘Every time we have looked into the sky with a new kind of detector, a new black box, we have found something which we did not expect.’

Weber never accepted that his detectors were not actually detecting gravitational waves and the mystery of just what they were detecting, if anything, has never been resolved. He continued working with his bars for the rest of his life, becoming isolated from the mainstream of gravitational wave detector research. His death was indirectly caused by his obsession with gravitational waves. In the winter of early 2000, a few months short of his eighty-first birthday, he was visiting his unmanned observatory on a hill in Maryland when he slipped on ice and broke several bones. He never recovered from complications caused by the injuries, and died eight months later on 30 September. He did not live to see the LIGO experiment become operational. And in an interview that is now in the archives of the American Institute of Physics, Charles Misner, a leading relativist, said, ‘The whole effort would never have been started if he hadn’t shown the world that you could take gravitational waves seriously. Before him, nobody did.’ As Bartusiak emphasises, ‘No one will take away [Weber’s] historic stature,’ and one measure of his place in history is that his first bar is now among the scientific artefacts at the Smithsonian Institution Washington, DC.

Lasers Shed Light

The direct line from Weber's work to the actual discovery of gravitational waves can be traced via Ron Drever, who was inspired by his visit to Weber's lab in Maryland to take up the challenge of detecting gravitational radiation. His own initial experiment, back in Glasgow, involved searching for vibrations in two masses, each weighing about 600 pounds; but he found nothing during a seven-month long run, and started thinking of other ways to carry out the search.

Weber was one of several people who, in the early 1960s, independently pondered the possibility of detecting gravitational waves using a technique known as interferometry. As the name suggests, this is based on the way two waves – in this case, two beams of light – interfere with one another. We have all seen what happens when a pebble is tossed into a calm pond. Smooth ripples spread out across the pond in all directions. But if two pebbles are tossed in to the pond at the same time, this produces two sets of ripples which interfere with one another to make a more complicated pattern. In some places the waves cancel out to leave the surface more or less flat; in other places the waves add together to make extra high ripples. Something similar happens in a classroom experiment with light to demonstrate its wave-like character. A beam of light is shone through two small holes in a screen and directed on to a second screen in darkened room (these 'screens' can just be pieces of cardboard). The light waves spreading out from each hole in the first screen interfere just like those ripples on a pond, producing a pattern of light and shade on the second screen – an interference pattern. What Weber, and a few others, realised is that using pure beams of laser light this interference could be used to measure very small changes in the distance between two objects; the kind of very small changes that would be produced by a gravitational wave squeezing and stretching the space between the two objects.

Weber never developed the idea himself, but he mentioned it to Robert Forward when they were building Weber's first detector, and Forward made a rough sketch to show how the experiment would work. A beam of laser light would be split into two beams, marching precisely in step with one another but travelling at right angles to each other along the two arms of the apparatus for exactly the same distance, before being reflected back along the same paths to merge again and make an interference pattern, which would be monitored by a detector. If the equipment was carefully aligned, the waves would interfere in such a way that they cancelled out, and the detector would detect nothing at all. But if a gravitational wave passed through the experiment, one arm would be squeezed and the other stretched, so that the two beams would get out of step, and a trace would appear in the detector and could even be displayed on a monitor screen as wiggly line, equivalent to the pattern of light and shade produced in the experiment with two holes.

People often ask, by the way, how the laser beams can detect the stretching and squeezing of space, since presumably they are also affected by the gravitational waves, being stretched and squeezed just as much as everything else. The answer is that this distortion of spacetime affects how long it takes for the light beams to get from one end of the experiment to the other, and light always travels at the same speed. So what the interferometer actually reveals is a time difference, not a space difference; that is automatically converted into the spatial equivalent in the detector.

At the end of the 1960s, Forward was working at the Hughes Research Laboratories in Malibu, California. He persuaded his employers to let him build a small interferometer, with the help of two of his colleagues, Gaylord Moss and Larry Miller. This was when it was still widely believed that Weber had actually detected gravitational waves, and the experiment was set up to be most sensitive to waves coming from the centre of the Milky Way, where Weber thought he had detected gravitational radiation. It was constructed more in hope than expectation of detecting anything, though, and was primarily intended to test the techniques that would be used in a full-size detector. The laser beams were directed along evacuated aluminium tubes just 2 metres long, with weights weighing only about a kilo each at the ends. It sat on a granite slab, which itself sat on a cushion of air jets, to minimise vibration, but it could only be run at night and weekends because there was so much extraneous ‘noise’ in the building during normal working hours. It ran, off and on, for 150 hours from early October to early December 1972, but found nothing. Forward dreamed of building a much bigger detector with arms at least a kilometre long, but the patience of his employers and the funds they were willing to provide had both run out, and he had to concentrate on more obviously practical work. But he described the Hughes interferometer to Ron Drever on a visit to Glasgow. Drever, disappointed by the failure of his own bars to detect anything, saw interferometry as the way ahead, and the Glasgow group began work on their own interferometric detector.

Drever's speciality was in the great British tradition of 'making do'. He was an expert at taking cheap bits of apparatus, sometimes scavenged from other experiments, and getting them to work. Although he had a limited budget, in Glasgow he was the undisputed leader of a small group, and could run things his own way. The result was that by 1976 the Glasgow team had a working interferometer. At the same time, a larger and better-funded group at the Max Planck Institute in Munich was making progress with their own interferometer experiment, in a friendly but nonetheless intense rivalry with the Glasgow group; the two teams would later join forces in a European collaboration. But the inspiration for the initial German effort came not from Robert Forward, but from another American, Rainer Weiss, who hit on the interferometer idea independently.

Weiss had been born in Germany in 1932, but his Jewish family fled to the United States just before the outbreak of the Second World War. In the late 1960s, he was working at MIT, where he was given the task of teaching the basic course on the general theory of relativity. As part of the course, he came up with the idea of a problem to set the students. He asked them to imagine three masses set up at three of the corners of a square, and to calculate how the distances between the masses would change as a gravitational wave passed through the arrangement. He then realised that what had started out as a classroom exercise could be turned into a practical experiment if laser beams were used to monitor the stretching and squeezing of space that was moving the masses in and out.

Like Forward, Weiss and his students built a small interferometer, with arms just 1.5 metres long, to test the idea. They were not seriously trying to detect gravitational waves at this point, but working out how to keep the laser beams perfectly aligned, the mirrors perfectly balanced, and so on. Weiss also analysed how the effects of unwanted vibrations (physicists call any unwanted interference noise) could be filtered out mathematically. He became convinced that the idea was viable, but that to achieve success would require a very big detector, with arms more than a kilometre long. As the next step towards such a detector, in 1973 Weiss put in a proposal to the US

National Science Foundation for long-term funding and development of his prototype experiment. But the application was rejected, and Weiss moved on to other work, becoming a key member of the team working on another big project, the Cosmic Background Explorer (COBE) satellite that was launched in 1989 to study the leftover radiation from the Big Bang. By the mid-1970s, all of the gravitational wave experimenters knew about each other – except for a Russian group working behind what was then still the Iron Curtain, with no easy communication with Western counterparts – but none of them had funding for the kind of major project that they each dreamed about. Things began to change in 1975, thanks to a boost for gravitational wave research provided not by any of the experimenters, but by theorist Kip Thorne.

Thorne had been a student of John Wheeler, and took over his mantle, becoming an expert on the general theory and the leading black hole theorist of his generation. In 1975, he was working at the California Institute of Technology (Caltech); in the wake of the discovery of the black hole associated with Cyg X-1 he was working with his students to improve the understanding of gravitational waves, and considering how best to get Caltech involved on the experimental side. At that time, Weiss was chairman of a NASA committee looking at possible relativity experiments to carry out in space, and invited Thorne to a meeting in Washington to give them some advice. The two of them hit it off immediately, and outside the formal committee meetings they stayed up all night discussing the possibilities. Before that meeting, Thorne had not been enthusiastic about interferometer experiments. Indeed, in a classic textbook, *Gravitation*, which he had co-authored with Wheeler and another relativist, Charles Misner, he had written that in this context laser interferometers 'have so low a sensitivity that they are of little experimental interest'. The all-night session with Weiss changed his mind and determined the path the Caltech experimenters would follow. With funding committed by Caltech, the NSF was persuaded to join in. Now all theorist Thorne needed to get the ball rolling was a top experimenter with experience of applying laser interferometry to gravitational wave research. The obvious choice was Drever, who now had a working instrument with arms 10 metres long (limited by the size of the converted particle accelerator lab where he was working). Drever was reluctant to move away from the British system, where he might be underfunded but could work in his own way, to the American system with more funding but going hand in hand with that more bureaucracy and less freedom. But in 1979 he was persuaded to spend half his time at Caltech to see if the project would work out. In 1984, he moved there full time as it gathered momentum.

Under Drever's expert guidance, an instrument with arms 40 metres long was constructed on the Caltech campus. But even this was only a 'proof of principle' project, a test bed for innovations (it is still in use as a test bed today). In the early 1980s, the Caltech instrument was capable of detecting a strain of one million-billionth (10^{-15}), equivalent to an atom moving across a distance equivalent to its own diameter. By the mid-1990s, various improvements had pushed this to 10^{-18}, capable of measuring a movement of the mirrors one-thousandth the size of an atom. But by then, the search for gravitational waves was moving into the big league, both physically and in financial terms.

LIGO and LISA

Weiss had never given up hope of building a ‘proper’ gravitational wave observatory, not a mere ‘detector’ but in effect a new kind of telescope, to monitor the Universe with gravitational waves the way radio telescopes monitor the Universe with radio waves. His experience with NASA had shown him that it wasn’t always necessary to make progress step by step, moving up (in this case) from detectors with arms 40 metres long to those with arms 400 metres long, then one a couple of kilometres across, and so on. When projects had very long lead times, like scientific satellites, it was better to make a commitment to the project and then develop the technology required while the project was being built. When the NSF, having previously turned down his application for funding for a modest project, started to put money into the Caltech project, Weiss decided to try again, and submitted a proposal for a much bigger experiment that would leap all the way from proof of principle to the actual detection of gravitational waves. By now back at MIT, at the end of the 1970s Weiss was given limited funding by the NSF to prepare a thorough feasibility study of what such a project might involve. The study was completed, with the help of colleagues at MIT, in 1983.

The proposal was nothing if not ambitious. It envisaged building a pair of identical detectors at widely separated locations, so that gravitational waves, which affect both detectors with a small time delay, could be distinguished from local disturbances affecting each individual detector. Each detector would have arms 10 km long, and the project would cost $70 million. There was a lot of resistance in the scientific community to putting so much money into a single project, which many people thought had little or no chance of success and would starve other proposals of funds. After a lot of deliberation, the NSF gave it the go-ahead in 1986, with the proviso that it had to be a joint project involving both MIT and the group at Caltech which was already being funded by the NSF. In practice, the size of the arms of the two detectors was reduced to 4 km, because of the limited sizes of the available sites, and construction work did not begin until the mid-1990s. The costs, though, got bigger instead of smaller – upwards of $1 billion. The detectors were built at Hanford, in Washington State, and Livingston, in Louisiana (the Louisiana site, in particular, was not ideal, but politics played a part in deciding where to spend the money). There were many trials and tribulations along the way, including the removal of Drever from the team as a result of the kind of clash that leaves musicians exiting from rock bands citing 'musical differences'. The gory details are described in the books by Marcia Bartusiak and Janna Levin. Here, though, it is time to cut to the chase and describe what the Laser Interferometer Gravitational-wave Observatory (LIGO), the most expensive project ever funded by the NSF, actually discovered, in September 2015.

While the detectors were being built and tested, the theorists, including Thorne, had ample time to refine their calculations of what kind of 'signal' they might detect. A prime candidate was the collision and merger of two black holes, in a binary system not unlike the binary pulsar but involving much more massive objects, as they spiralled together. Simulations of the way such a pair of black holes would spiral together and merge, according to the general theory of relativity, were carried out for a variety of different black hole masses. The distinctive feature of such a merger, in gravitational wave terms, is a 'chirp', a signal in which the ripples in spacetime get higher in pitch (shorter wavelengths) as the black holes spiral closer together, then ends abruptly as the merger is completed. If you could hear it, it would be a bit like the sound you get by running your hand rapidly along the keys of a piano from left to right, Jerry Lee Lewis style. So the experimenters knew exactly what they were looking for. They also had a self-imposed deadline. Einstein had completed his general theory in November 1915 and immediately presented it to the Prussian Academy of Sciences; it was formally published early in 1916, the year he also found that the theory predicts the existence of gravitational waves. The centenary of that publication was the target the LIGO team had set themselves to be up and running, and hopefully finding something.

With that in mind, the first proper science run of the detectors was planned for September 2015, and the final engineering test run, slightly behind schedule, was being carried out in the small hours on Monday, 14 September. During those tests, there was an interval of just under an hour when nothing was being done to interfere with the machines, and they were left locked in observing mode, each with a lone operator keeping an eye on things. At 2.50 a.m. local time in Hanford, and 4.50 a.m. local time in Livingston, almost simultaneously, each detector recorded a chirp lasting 200 milliseconds. There was no drama – no flashing lights or ringing bells, and nothing for human eyes to see on detector screens. But the computerised system went into action, recording the data and sending messages to a select band of researchers around the world to let them know what had happened. When the scientists got in to work at their usual time, their first reaction was that their colleagues had deliberately injected a false pulse into the system as a test – a so-called blind injection, to find out if anyone noticed it and see how the systems reacted. It wasn't. The detectors had picked up a gravitational wave signal far stronger than the team had expected, and far more quickly than they had anticipated.

Because the pulse was seen in two identical detectors, one on each side of the North American continent, with a delay of just 6.9 milliseconds between its arrival at the first detector and its arrival at the second detector, this confirmed that it was real, and that it travelled at the speed of light.

The exact pattern of ripples in the pulse matches the predictions for the collision and merger of two black holes, one with about 29 times the mass of our Sun and the other with a mass of about 36 times the mass of our Sun; the waves detected were produced during the last ten orbits of the pair around each other, with a final 'ring down' as they formed a single oscillating black hole with a mass of about 62 times the mass of our Sun. As those numbers indicate, in the process about three times the mass of our Sun was converted into energy in the form of gravitational waves, in line with Einstein's famous equation $E = mc^2$. This enormous outburst of energy, the most powerful event recorded by human observers, was equivalent to 10^{23} times (a hundred billion trillion times) the luminosity of our Sun. The distance to the source was estimated as nearly a billion and a half light years, and it was able to shake detectors on Earth by a tiny amount; the change in the length of each 4 km-long arm was roughly one ten-thousandth of the width of a proton, a strain of 10^{-21}. Because the wave hit the Louisiana detector first, the astronomers had a rough idea of the direction it had come from, somewhere on the southern sky. But there is no sign of any cosmic catastrophe visible in that region at optical or radio, or any other electromagnetic, wavelengths. LIGO really has opened up a new window on the Universe. Because the implications of the discovery are so profound, the team checked and double-checked their data before sending off their scientific paper for publication. So the news did not break officially until 11 February 2016 – ruling out, as I mentioned, any chance of the Nobel Prize being awarded to Drever, Weiss and Thorne that year. (Sadly, Drever died a few months later, and Nobel Prizes are never awarded posthumously.) Even before the paper was published, however, LIGO had detected a second black hole merger. This event shook the detectors on Christmas Day 2015 (at least, it was Christmas Day in the United States. Some reports give the time of the event on the system used by astronomers, which is essentially the same as GMT; that puts it in the early hours of Boxing Day.)

It was much smaller than the first event, producing the characteristic chirp caused by the merger of black holes with 14 and 8 times the mass of the Sun, combining to make a black hole with a mass 21 times that of the Sun, with about one solar mass of matter being converted into energy. But although the signal was weaker, the detectors recorded the final couple of dozen orbits of the binary, taking a full second to complete. This observation was doubly significant. It showed that the detectors are sensitive to less energetic events, and it justified the name ‘observatory’, proving that the first detection was not a one-off fluke. Already, astronomers have gone beyond detecting gravitational waves to observing gravitational waves regularly (there have since been more detections) and using them to probe the Universe, addressing questions like how many black holes there are out there, and where they came from.

Where do we go from here? LIGO itself is being upgraded and will be running late in 2017 with a sensitivity that should allow it to detect neutron star mergers taking place in our own Milky Way Galaxy, as well as roughly one black hole pair merger per day. Meanwhile, a similar detector has become operational in Italy, and another is being completed in India. With several detectors in operation, the timing of the delays in the arrival of the waves at different places will make it possible to pin down the sources of the waves more accurately on the sky, perhaps making it possible to identify optical or radio counterparts to these events. The greatest hope of the observers is that they will find something completely unexpected, like the accidental discovery of Sco X-1 in 1962. But in the longer term, the most exciting possibility is a planned gravitational wave observatory in space, known as LISA (Laser Interferometer Space Antenna).

LISA has had a chequered history, and is not yet guaranteed to get off the ground; but the success of LIGO has given its prospects a boost. The idea of a laser-based gravitational wave observatory in space has been around since the early 1970s, and was discussed by pioneers including Drever, Forward and Thorne. But it would be two decades before technology became sufficiently advanced for the European Space Agency (ESA) to take the idea seriously and accept it as a future mission in 1993. From the outset, ESA expected that NASA would join the project, and this duly happened in 1997.

As originally envisaged, LISA would have involved three separate spacecraft placed at the corners of a triangle with sides (arms) five million kilometres long, following the Earth in its orbit but 50 million kilometres behind us. Each satellite would contain two perfect cubes of polished platinum-gold 1.5 inches wide, floating freely inside the spacecraft so that they were falling around the Sun only under the influence of gravity. Nothing inside the satellite touches the test masses, except for laser light. Each satellite would carry two lasers, linking it to each of the other two satellites. But with such long distances involved, instead of the lasers simply being reflected back to their sources, they would have to be amplified and then returned to make an effective interferometer system. And in any one second, the satellite must not move relative to its freely-falling test masses by more than a few nanometres – the thickness of a few atoms (one nanometer is a billionth of a metre, 10^{-9} m). This requires the development of the most gentle thrusters ever built. If it all worked, the three-cornered observatory would stay in orbit for several years, monitoring the ripples in spacetime.

But this would not be a competitor to LIGO and other ground-based instruments. A detector as large as LISA would be sensitive to very long-wave disturbances, which have been likened to long ocean swells, rather than the ripples monitored by LIGO. Such waves would take hundreds of seconds, peak to peak, to pass through the triangular array, stretching and squeezing the space between the satellites. This means that whereas LIGO is sensitive to events involving black holes with a few, or a few dozen, times the mass of the Sun, LISA would be sensitive to events involving black holes with hundreds or millions of times the mass of the Sun. This is particularly exciting because this kind of black hole that is thought to lie at the centres of galaxies. LISA would also be able to detect gravitational waves from neutron star binaries where the two stars are still far apart, whereas LIGO should be able to detect them in the final stages of their in spiralling before they collide and merge, perhaps to make a black hole.

The whole project was technologically advanced, and very expensive.

Too expensive, as it turned out. In 2011, NASA was being forced to make cuts to many projects, in large measure because of cost overruns on its flagship project, the James Webb Space Telescope, the successor to the Hubble telescope. It pulled out of LISA, which might have killed the project. Apart from the financial implications, the project had depended on the availability of large launch vehicles supplied by NASA; the only alternative would be the venerable, reliable but much smaller Russian Soyuz launchers. Could a smaller, cheaper version of LISA be designed to fit both constraints? Reducing the lengths of the sides of the triangle to 1 million kilometres would help, because less fuel would be needed to place the satellites in place (for comparison, the diameter of the Moon's orbit is 770,000 km). Such a cut-down detector would still be able to monitor waves from neutron star binaries and from black holes with 10,000 to 100,000 times the mass of the Sun; mergers of such black holes are thought to be the process which builds up the even larger black holes at the centres of galaxies. A more drastic idea was to cut out one side of the triangle, leaving a V-shaped configuration with just two laser systems. Such thinking led to a revised proposal, formally known as the Evolved Laser Interferometer Space Antenna, or eLISA, (nobody except the bureaucrats uses the name eLISA, though; all the astronomers I know still call it LISA), which ESA selected in 2013 for a planned launch in the mid-2030s, probably still with a three-sided configuration, with sides 1 million km long, with a planned lifetime in orbit of a couple of years.

By then, a test bed for the LISA technology, LISA Pathfinder, was already being prepared, having been planned and developed for nearly ten years. It was launched in December 2015, after the discovery of gravitational waves by LIGO but before the discovery was announced, and exceeded all expectations, ensuring the green light for the full project.

Although primarily an ESA mission, NASA made a contribution to LISA Pathfinder to test some of the technology required for the laser rangefinding and other engineering of the project. This mission alone cost more than $600 million, and the spacecraft was about the size of a small family car. After being steered to its orbit, at a particularly stable point relative to the Earth known as Lagrange 1, in February 2016 the two test cubes, each 46 mm on a side and with a mass of two kilos, were gently released from the constraints that had held them tight during launch. This tricky procedure involved first gently pulling away the main constraints, leaving the test masses held on the points of needle-like fingers, then ever so delicately pulling those fingers away. The test masses were left inside two vacuum chambers (the vacuum had to be better than the vacuum of ‘empty space’) 38 cm apart. The cubes were then falling freely through space, while the spacecraft was essentially flown around them, its orbit being precisely maintained by the tiny thrusters – the most delicate rockets ever flown – to compensate for outside influences such as drag caused by the impact of atoms and molecules from the imperfect vacuum outside. The trick was to do this and keep the walls of the spacecraft from bumping into the cubes.

The distance between the test masses was monitored with a laser, a mini-interferometer which showed that the two cubes stayed motionless with respect to one another with amazing precision. The accuracy called for by the Pathfinder plan was reached on the first day of operation, and over the rest of the experiment, which lasted for six months in all, they were kept a precise distance apart to five times the accuracy originally hoped for. In terms of numbers, the aim was to reach an accuracy of a picometre maintained for several hours – a millionth of a millionth of a metre, 10^{-12} m. The final accuracy achieved was on the femtometre scale (a femtometre is 10^{-15} m, so a few hundred femtometres is a fifth of a picometre), and the cubes were kept stable to better than the accuracy that will be required for the full LISA observatory (or eLISA if you are a bureaucrat) for fifty-five days. As they say in the movies, ‘we have the technology’ to make it all work. What we need now is funding – a couple of billion dollars – and a launch vehicle. Watch this space.

Further Reading

Marcia Bartusiak, *Einstein's Unfinished Symphony* (Joseph Henry Press, 2000) The clearest and most comprehensive account of the history of gravitational wave research. The only flaw is that it was published before the LIGO detection, but a new edition is promised for 2017.

John Gribbin, *Einstein's Masterwork* (Icon, 2015)
My take on the general theory of relativity.

Janna Levin, *Black Hole Blues* (Bodley Head, 2016)
A gossipy and entertaining account from a writer who had privileged access to the LIGO team, published after the first detection of gravitational waves.

ENDS