The General Theory of Relativity

The Special Theory of Relativity, discussed in subunit 7.1, is restricted in its application to inertial frames of reference. The first postulate says that the laws of physics have the same mathematical form in all inertial frames of reference, an inertial frame being one in which the laws have their simplest mathematical form. Once an inertial frame is found, all other frames moving in a straight line with a constant speed are also inertial frames. This creates a special class of observers and excludes from the theory any frame of reference accelerating (changing either its speed or direction of motion) with respect to an inertial observer. A frame of reference in which there is a gravitational field is also a non-inertial frame and is excluded from the special theory. The laws of physics have a more complicated mathematical form when gravity is present. Einstein's physical intuition told him that the principle of relativity could be extended to include non-inertial frames of reference, and by doing this, gravity would emerge as a natural property of space-time.

Einstein set out almost immediately to develop this more general theory that would include non-inertial as well as inertial frames. The task was formidable, requiring a degree of mathematical sophistication previously unknown in physics. By late 1915, the task was complete.

It is significant to note that Einstein was led to the more general theory not by the need to explain some experimental data, but rather the "need" to conform to certain philosophical principles. In his own words,

> "Our experience hitherto justifies us in believing that nature is the realisation of the simplest conceivable mathematical ideas. I am convinced that we can discover by means of purely mathematical constructions the concepts and the laws connecting them with each other, which furnish the key to the understanding of natural phenomena. Experience may suggest the appropriate mathematical concepts, but they certainly cannot be deduced from it. Experience remains, of course, the sole criterion of the physical utility of a mathematical construction. But the creative principle resides in mathematics. In a certain sense, therefore, I hold it true that pure thought can grasp reality, as the ancients dreamed."

This approach may be criticized as being unscientific (and it is in the case of most ordinary people), but in the case of Einstein, it produced a physical theory of incredible beauty and power.

The Principle of Equivalence

The concept of mass is familiar. In most people's minds, mass is associated with weight. More massive objects weigh more. This property of mass is called gravitational mass, because it determines the force that gravity will exert on the object. There is another, seemingly unrelated, property of mass. The more massive an object is, the harder it is to change either the direction or the speed of its motion: that is, the more difficult it is to accelerate the object. This property of mass is called inertial mass. In classical physics, inertial effects are unrelated to gravitational effects. In developing his theories of motion and gravity, Newton assumed that they were the same property, although there was no theoretical justification for this assumption. It just seemed to work. By the end of the nineteenth century, it had been empirically shown that, to a high degree of accuracy, inertial and gravitational mass can be considered numerically equal, although there was still no theoretical explanation for this.

Having two completely independent properties that are accidentally equal to each other is, philosophically, a very unsatisfactory situation. Einstein reasoned that there must be some underlying physical significance to this equality which would constitute a single interpretation for mass.

As with the special theory, Einstein used thought experiments to guide his thinking. Imagine a person in a small, windowless compartment with two objects of different mass. Suppose this compartment is located on the surface of the earth. If the two objects are dropped simultaneously from the same height, they will hit the floor at the same time. Galileo was the first to show this and to measure the acceleration due to gravity as 9.8 m/sec^2. That is, near the surface of the earth, a free-falling object will change its speed at a rate of 9.8 m/sec each second.

Suppose now that the compartment is in space far removed from all massive objects. There will be no gravitational forces acting on the compartment and its contents. If there are no forces acting on the compartment, the person and the two objects will float around in the compartment much as you see the astronauts doing when in orbit around the earth. This is often referred to as weightlessness, but that is not strictly correct. The astronauts are not weightless; the earth still pulls them and the space craft downward. That weight is what keeps the space craft in orbit rather than flying off into space. However, with the rockets turned off, the space craft free falls around the earth, and this is equivalent to being weightless.

Now, imagine that there are rockets attached to the bottom of the compartment, causing it to accelerate. The inertial mass of the person, with its tendency to maintain its state of motion, will resist this change in velocity. The person will be pressed against the floor of the compartment just as the driver of a car is pressed against the back of the seat as a car accelerates. With effort, the person will be able to stand. If the two objects are dropped as before, their inertial mass will tend to keep them in a constant state of motion, and the floor of the compartment will accelerate up and overtake them at an ever increasing speed.

In the frame of reference of the observer, the objects will “fall” to the floor. In fact, if the acceleration of the compartment is 9.8 m/sec per second, they will “fall” to the floor exactly as they would if the compartment was at rest on the surface of the earth. As Einstein thought about this, he realized that there is no possible measurement that the person in the compartment could make that would distinguish between the two situations. In 1911, based only on his thought experiments, Einstein made the following bold statement: "It is impossible to distinguish, by any experiment whatsoever, between the effects of acceleration and the effects of gravity." This statement is known as the principle of equivalence.[1]

The thought experiment above is a mechanical one. What about experiments involving electromagnetic phenomena? Again, as was the case in 1905, Einstein turned to the propagation of light. Suppose that there is a source of light in an inertial frame of reference and that a beam of light from the source enters the compartment horizontally from the side. If the compartment is at rest with respect to the light source, it is also an inertial frame. The light beam will enter the compartment, travel horizontally across the compartment, and strike the wall the same distance above the floor as the place it entered. By considering this situation in the frame of reference of the accelerating compartment in outer space (a non-inertial frame of reference), Einstein was able to make an important prediction.

In an inertial frame, the speed of light is an absolute constant. Thus, the beam of light in its inertial frame travels at a constant speed as it passes through the compartment. However, the vertical speed of the compartment constantly increases. Thus, in fixed time intervals, the beam of light will always travel the same horizontal distance, while the compartment will travel increasingly greater vertical distances.

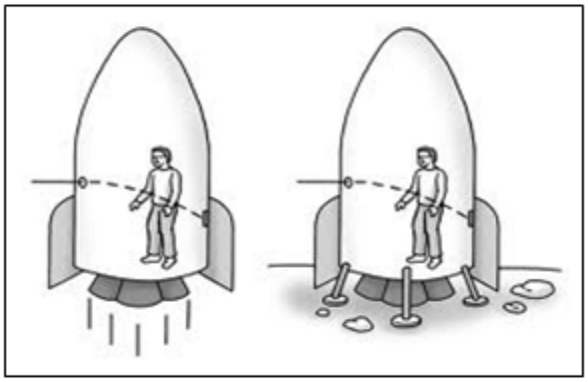
In this case, the path of the beam of light is not a straight line in the frame of reference of the compartment. It will be bent toward the floor. If an accelerating frame of reference is in no way distinguishable from a frame of reference in a gravitational field, then this thought experiments predicts that a beam of light will not be a straight line in a gravitational field. Einstein concluded that the path of light is altered by gravity such that a light beam passing near a massive object ought to travel as if attracted toward the object.

Calculations showed that the amount of bending produced even by an object as massive as the sun is extremely small. However, in a 1911 paper, Einstein proposed an experiment that should in principle allow a measurement to be made. For reasons discussed later, this experiment was, fortunately, not carried out until 1919.

---

[1]Footnote. The principle of equivalence only applies in a small region of space. Otherwise non-uniformities in the gravitational field can be distinguished from the uniform effects produced by acceleration. The principle of equivalence should not be confused with the equivalence of mass and energy, which is an entirely different matter.

The Principle of Equivalence



A beam of light is bent in exactly the same way in a frame of reference that is accelerating and in one in a gravitational field. The inability to distinguish between the two experimentally is known as the principle of equivalence.

The General Theory of Relativity

Einstein's development of the General Theory of Relativity was strikingly similar to his development of the Special Theory of Relativity. In the special theory, Einstein took the law of propagation of light and the principle of relativity (for inertial frames of reference), which appeared to be incompatible and showed that they were both true. In the case of the general theory, he took the principle of equivalence and the general principle of relativity (that the laws of physics are the same in all frames of reference, both inertial and non-inertial), which appeared to be incompatible and showed that they were both true.

In the previous section, it was shown that the principle of equivalence led to the fact that while light in an inertial frame of reference travels in a straight line, light in an accelerating frame (or in a gravitational field) travels in a curved path. It appears that the laws governing the motion of light in an inertial and in a non-inertial frame are different, violating the general principle of relativity. That Einstein was able to resolve this paradox with the General Theory of Relativity, in spite of unimaginable mathematical and conceptual difficulties, stands as a monument to the human intellect. As was the case with the special theory, the solution lies in our concept of space-time.

The main problem Einstein had with developing his new theory was mathematical. Although by ordinary standards, Einstein might be considered a mathematical genius, by the standards of theoretical physicists, Einstein was not exceptional. It was his unprecedented physical intuition (with the possible exception of Isaac Newton) rather than his mathematical ability that made him the greatest physicist of his time. It was only in the process of work on his theory that Einstein gradually acquired the mathematical techniques with which to express the theory. In late 1912, he wrote to a friend:

> "I occupy myself exclusively with the problem of gravitation and now believe that I will overcome all difficulties with the help of a friendly mathematician here. But one thing is certain: that in all my life I have never before labored at all as hard, and that I have become imbued with a great respect for mathematics, the subtle parts of which, in my innocence, I had till now regarded as pure luxury. Compared with this problem, the original theory of relativity is child's play."

The "here" was his alma mater, Zurich Polytechnic Institute, where he had just returned as professor of physics, and the "friendly mathematician" was Marcel Gossmann, an old school friend. The necessary mathematical technique was tensor calculus, Grossmann's specialty.

In addition to the principle of equivalence and the general principle of relativity, Einstein put an additional constraint on the theory, one dictated by aesthetic values, which he held to be of paramount importance in physics. Out of literally thousands of formalisms provided by the tensor calculus consistent with the principle of equivalence and the general principle of relativity, Einstein insisted that only the mathematically simplest formalism would provide a correct description of nature.

In 1914, Einstein left Zurich for Berlin. There, late in 1915, after years of almost constant effort, Einstein arrived at a formalism that seemed to satisfy all his requirements. As a test, Einstein used the theory to calculate the orbit of Mercury. The orbit calculated using the classical Newtonian theory of gravity was very, very slightly, though undeniably, inconsistent with the observed orbit. The difference between the two was an incredibly small 43 seconds of arc per century in the precession rate of the orbit.

However, using his new theory, Einstein's calculated orbit for Mercury matched the observed orbit perfectly; the additional 43 seconds of arc per century came naturally and of necessity from the theory. It was immediately clear that Einstein's theory is a more accurate description of gravity than Newton's, a theory that for centuries had been assumed to be absolutely correct.

The General Theory of Relativity is, as Einstein had intended from the beginning, a theory of gravity. However, the general theory did not simply produce a new, more general force law. It changed in a very fundamental way our concept of gravity. According to the theory, the effects of gravity are not the result of a force being exerted on an object but rather are the result of the natural, inertial motion of the object through space-time, the properties of which are determined by the presence of other massive objects.

In the special theory, space and time are interwoven and separate models of each cannot be constructed. In the general theory, space-time and matter lose their independent meaning. Space-time and matter are different aspects of a single unity; each is meaningless in the absence of the other.

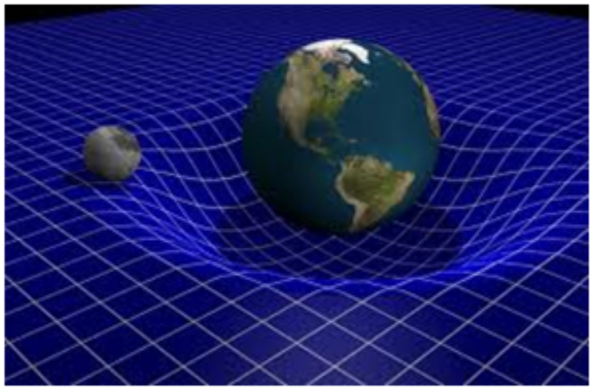
In addition to providing a more accurate description of the effects of gravity, the general theory also eliminated one of the most disturbing aspects of the classical theory—action-at-a-distance. The sun does not pull on the earth, rather the sun changes the properties of the space-time in its vicinity. In effect, the sun curves space-time in a way that inertial motion through it is no longer a straight line.

This is much like an object moving on a curved two-dimensional surface. Imagine a horizontal, flat rubber sheet. A bowling ball placed on it will depress the sheet, creating a curved surface surrounding the ball. If a marble is rolled across the rubber surface, it will not travel in a straight line. The marble will be deflected by the curved surface with the amount of the deflection being greatest where the curvature is greatest. In fact, the speed of the marble can be such that the marble will orbit the bowling ball, much as the earth orbits the sun. The bowling ball does not exert a force on the marble to produce the orbit, rather the curvature of the surface produces the orbit.

The amount of the curvature of the rubber sheet depends on the distance from the ball. In this analogy, the amount of the curvature of the rubber sheet models the strength of the gravitational field. Just as is the case with the strength of the gravitational field, the amount of curvature of the rubber sheet surrounding the bowling ball decreases with distance.

This analogy is not exact, as it explains the motion in terms of a curved two-dimensional surface, whereas gravity is a consequence of curved four-dimensional space-time. However, it does provide a visual way to represent the nature of gravity in the General Theory of Relativity.

Curvature of Space-Time



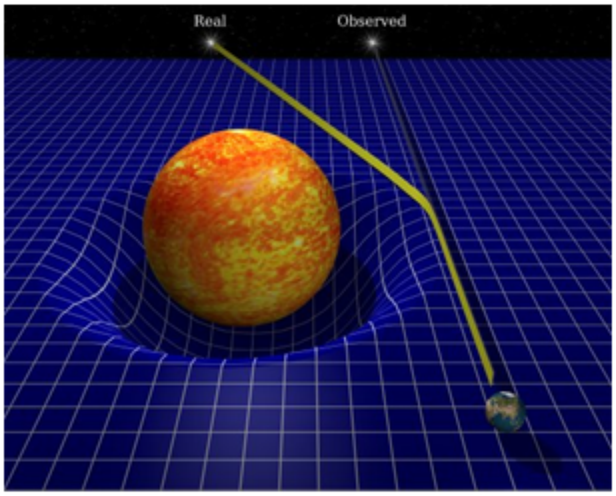
In the general theory, gravity is not a force acting at a distance. Rather its effect on motion is the result of a massive object, such as the earth in this example, curving the space-time in its vicinity. The moon orbits the earth as a result of its inertial motion through the curved space-time, not because the earth is pulling on it.

Evidence for the General Theory of Relativity

The calculation of the correct motion for the planet Mercury was a tremendous success for the general theory. However, physical theories are judged primarily on predictions of new, unsuspected physical phenomenon that are subject to experimental or observational verification. The Mercury result, though completely unforced, was not a prediction but an explanation of a previously known fact.

The first crucial test of Einstein's theory was made during the total eclipse of the sun in May, 1919. Einstein had proposed an experiment to measure the amount of bending of light as it passed near the sun. During a total eclipse of the sun, the sky is as dark as at night, and the stars can be seen. A photograph of the stars taken during a total eclipse can be compared to an earlier photograph of the same region of sky taken at night. Light rays passing near the eclipsed sun are bent. For these stars, their relative positions with respect to the other stars will be slightly different on the two photographs.

Deflection of Light by a Massive Body



The deflection causes the star to appear in a different point in the sky. The general theory predicted the shift in position that would result.

For the star in the figure, the photograph taken during the eclipse would show a different location in the sky compared to a photograph taken at some other time. The difference between the apparent and the true positions of the star is a direct measure of the amount of bending produced by the gravitational effect of the sun.

In 1911, Einstein predicted the bending of light in a gravitational field well before his theory was complete. He made this prediction based on the principle of equivalence and Newton's theory of gravity. The calculation indicated that a light beam passing very near the surface of the sun would be deflected through an angle of 0.87 seconds of arc or 0.00024 degrees. Although this angle was extremely small, there was a possibility that it could be measured. In 1914, the German astronomer Erwin Finlay-Freundlich, set off for Russia to observe the total eclipse of the sun and try to verify Einstein's prediction. However, he was prevented from doing so by the outbreak of the First World War. Einstein wrote the following in a letter to a friend:

> "Europe, in her insanity, has started something unbelievable. In such times one realizes to what a sad species of animal one belongs. I quietly pursue my peaceful studies and contemplations and feel only pity and disgust. My dear astronomer Freundlich will become a prisoner of war in Russia instead of being able there to observe the eclipse of the sun. I am worried about him."

In fact, it was fortunate that Freundlich was not able to make his measurements. By the end of 1915, it was clear that his 1911 prediction was in error. In 1911, he had assumed that Newton's gravitational theory was appropriate for the calculation, but his general theory indicated that this was not the case. A new calculation based on the general theory produced a value of 1.75 seconds of arc, about twice the earlier prediction.

There was, of course, no chance to test the eclipse prediction until the war ended. However, as early as 1917, Sir Arthur Eddington, a British physicist, began preparing for two 1919 expeditions. Eddington was one of the first to appreciate the significance of the general theory, but his efforts to organize the test expeditions were motivated by more than just scientific curiosity. Eddington was a Quaker, and like Einstein, was profoundly disturbed by the war. He felt that if a British expedition verified the work of a German theoretical physicist, this would help heal the wounds of war and, in particular, would reestablish scientific relations between the warring nations.

The results confirmed Einstein's prediction. 1919 was the last year of Einstein's private life. The announcement of the verification of his theory to a war-weary and heartsick people made him a world-wide hero. He became the personification of intelligence, an identification that has survived in spite of the fact that Einstein has been dead for more than a half-century. This fame, which he neither sought nor enjoyed, would later cause him to refer to his years of obscurity in the Bern Patent Office as the happiest of his life.

The General Theory of Relativity Today

In the decades following its publication, the number of observable effects that distinguished Einstein's General Theory of Relativity from the much simpler Newtonian theory was small, and the magnitude of the differences between the predictions of the two theories was almost negligible. For these and other reasons, interest in the general theory soon all but disappeared—almost, but not quite. During the twenties and thirties, a handful of theoretical physicists and cosmologists applied the general theory and arrived at results which staggered the imagination, results so inconceivable and incomprehensible that they were generally ignored.

This attitude toward the General Theory of Relativity changed rather dramatically in the 1960s due primarily to developments in the field of astronomy. The discoveries of such exotic objects as quasars and pulsars suddenly made the bizarre predictions of Einstein's theory seem less unreasonable. Today, the General Theory of Relativity is again at the forefront of physics. The two most important applications of the theory are the physics of gravitationally collapsed objects, black holes being the most extreme example, and cosmology, the science of the Universe as a whole.

A black hole is a region of space-time where gravity is so strong that anything that enters the region, even light, will be trapped there. The possibility of the existence of a black hole was recognized as a direct prediction of the general theory almost immediately after the formulation of the theory. However, it was not until 1939 that Robert Oppenheimer and one of his students, Hartland Snyder, suggested a possible mechanism whereby one might actually form.

When all possible fuels of a star are exhausted, it will begin to collapse gravitationally under its own weight. Using the general theory, Oppenheimer and Snyder were able to show that if the mass of the collapsing star exceeded a certain value, today believed to be about three times the mass of our sun, no known force could prevent complete collapse. The star will collapse to smaller and smaller volumes, collapse to the point where electrons, protons, and neutrons are crushed out of existence—crushed to an object of infinite density surrounded by a volume of space, where gravity will prevent the emission of light. It is little wonder that at first physicists were reluctant to accept this. Arthur Eddington characterized the idea as absurd.

Today, the existence of black holes is a virtual certainty. Black holes from 3 to 10 or so solar masses have been identified in orbit around ordinary stars. A supermassive black hole, containing 2.2 million solar masses, has been found in the center of our galaxy. Other supermassive black holes have been detected in the centers of other galaxies. Modern cosmology, one of the greatest intellectual adventures ever undertaken, owes its existence to Einstein's General Theory of Relativity. In 1916, soon after the calculation of Mercury's orbit, Einstein applied his theory to the universe as a whole and got an unwelcome result. The theory indicated that the universe could not be static. The theory clearly predicted that the amount of space in the universe must be either increasing or decreasing. Einstein did not believe that this was true, and the other physicists and astronomers he consulted assured him that it was not. Everyone was sure that the universe was static and unchanging on the large scale.

For once in his life, Einstein lost faith in the basic simplicity of nature and added an ad hoc term to his theory, which he called the cosmological constant. Though it marred the beauty and simplicity of the theory, Einstein believed that it would eliminate the offending prediction of expanding or contracting space. The effect of the cosmological constant is to provide a universal repulsive force that Einstein thought could balance the attractive force of gravity and allow for a static universe.

In the 1920s, two theoretical physicists independently showed that Einstein's "fix" of his theory did not work. Not only did the original theory suggest that space must be expanding or contracting, but the new version with the cosmological constant made the same prediction. Hardly anyone at the time studied the general theory, and no one paid much attention to these papers, including Einstein himself. One of the involved physicists, the Belgium priest, George Lemaitre, was so discouraged by the lack of interest in his paper that he switched his research to another field.

In 1929, the American astronomer, Edwin Hubble, published a law stating that the radiation received from distant galaxies is stretched out, i.e. has longer wavelengths, than radiation received from local sources. Furthermore, the amount of the stretching, called redshift by astronomers, is directly proportional to the distance to the galaxy.

Hubble knew that this was a very significant result but had no idea what was the cause of it.

When Arthur Eddington heard of this result, he immediately knew the explanation. Eddington was among the few physicists who actually understood the details of the general theory. He was also one of the few who were aware of the papers, showing that regardless of the version of the theory used, the general theory required either an expanding or a contracting universe. Either was consistent with the general theory. As to which described our actual universe, it was an empirical question. Eddington immediately recognized Hubble's Law as empirical evidence that space is expanding. When Einstein heard of this, he called his inclusion of the cosmological constant the "greatest blunder of my scientific career." We will later see that this may not be the case.

Eddington concluded space was expanding, because expanding space would stretch out the wavelengths of light from distant galaxies just as Hubble had observed. Further, the light from the more distant galaxies would spend more time traveling through expanding space and thus be stretched more, making the amount of redshift directly proportional to the distance. Had space been contracting, the wavelengths would be compressed, or blueshifted rather than redshifted.

When George Lemaitre learned of these developments, he returned his interest to cosmology. He was the first to think scientifically about what the universe must have been like in the past if space is expanding. He concluded that if time could be run backwards, the universe would become increasingly dense. Using the general theory, it was clear that this increasing density would become infinite at some finite time in the past, i.e. that the universe must have had a beginning in time.

The prevailing belief among scientists at the time was that the universe was infinitely old. Thus, the prediction by a Catholic priest that the universe had a finite age, (he had probably believed this from childhood) was not taken very seriously. Lemaitre's theory had an additional strike against it. Almost nothing was known about nuclear physics at the time, so Lemaitre was unable to use his theory to make testable predictions regarding the present universe.

The idea was taken up again in the late 1940s and early 1950s by the Russian-American physicist, George Gamow. By that time, considerable progress had been made in the area of nuclear physics, and Gamow was able to make two significant, testable predictions. The first of these was that the very early universe could not have produced any appreciable amount of chemical elements more massive than helium. Thus, the early universe must have nearly pure hydrogen and helium. The second was that the entire universe would be filled with thermal radiation at a temperature of a few degrees above absolute zero; the cosmic background radiation. Gamow's theory soon became known as the Big Bang theory.

By the late 1950s, it had become clear that the elements heavier than helium are actually created in the interiors of massive stars and that the explosive deaths of these stars distributed the heavier elements throughout the galaxy, making them available to later generations of stars, including our sun. Evidence accumulated, reinforcing that in the past the galaxy was nearly pure hydrogen and helium.

The clinching piece of evidence in favor of the Big Bang was the discovery in 1964 of the predicted cosmic background radiation. Arno Penzias and Robert Wilson, two radio astronomers working for Bell Laboratories, detected the radiation using a radio telescope built for trans-Atlantic telephone conversations. The pendulum of scientific opinion immediately swung from the Steady State theory, with an infinitely old universe, to the Big Bang and a finite age for the universe. Our most recent data from cosmic background radiation studies indicates that the universe began 13.7 billion years ago.

It is now universally accepted that space is expanding. It is also universally accepted that gravity, as understood by the General Theory of Relativity, requires that the rate of expansion must be decreasing. In the mid-1990s, two independent research teams set out to measure the rate of slow down. By 1998, their results were in. Much to their (and everyone else's) amazement, both determined that the expansion rate was actually increasing. The only possible explanation for this is that, in addition to the attraction of gravity, the universe contains a global repulsive force, and at the present time, the effect of this repulsive force is greater than that of gravity. The nature of this force is unknown, and it is just referred to as dark energy. Einstein's cosmological constant of 1916 seems to be in accord with what has been observed so far of the changing expansion rate of the universe. Perhaps its inclusion was not a "blunder" after all.