Gravitation

Hermann Bondi

Newton's theory of gravitation is presented in modern form. It is shown how, combining it with special relativity, one is led to Einstein's theory of gravitation ('general relativity'). Its consequences for orbits, black holes, gravitational energy and gravitational waves are briefly described.

Galileo's principle

Modern science may be said to have begun with gravitation, yet this subject is still an imperfectly understood fairly separate branch of physics. This article tries to take the reader through a series of considerations illuminating different aspects of it.

The essential starting point, which brings out the peculiarity of gravitation, is Galileo's statement that all bodies fall equally fast. This is still the essence of the subject. At first sight one might think that Galileo's principle, as I like to call it, would simplify matters, but the opposite is the case, since we cannot compare a body on which gravity acts with one that is immune to such a force, there being no 'immune' materials. In a sense Galileo's principle was already implicit to the acceptability of the Copernican scheme because, if different substances on the Earth reacted differently to the Sun's gravitational pull, we would see a sorting of matter which is not in fact observed.

Galileo inferred his principle from experiments which were as good as the technology of his day, more than 3½ centuries ago, allowed. The principle has been tested since with more modern means. Early in this century Eotvös in Hungary carried out a most ingenious experiment which showed that deviations from Galileo's principle could not exceed one part in 10^7, and in the seventies R. H. Dicke at Princeton reduced this to one part in 10^{11}. Thus we know that the principle is satisfied to more precision than we can claim for most other physical principles. Yet nagging doubts remain. Gravitation is a remarkably weak force, weak compared with others that we know. A typical case of another force is the electrostatic attraction between charged bodies, the force on which all atomic structure depends. In the simplest case, the hydrogen atom, there is one (positively charged) proton and one negatively charged electron. The electrostatic attraction between them can be effectively measured, while the tiny gravitational attraction can only be inferred. The ratio between the two forces (which is independent of distance) is a few times 10^{39}, a truly enormous number!

Hermann Bondi is a Fellow of Churchill College, Cambridge, CB3 0DS, UK.

Thus the second characteristic of gravitation is its extreme weakness, a weakness so great that one is surprised that gravitation is at all measurable. In fact it is an immediately noticeable force in our surroundings. How does this come about?

Gravitation's third characteristic is that it is always of the same sign. A positively charged particle on the surface of the Earth is attracted electrostatically by all the negatively charged particles in the Earth (essentially electrons) and repelled by all the positively charged ones (protons), but gravitationally *all* particles attract it.

The electrostatic attraction and repulsion are so perfectly balanced that the resultant force effectively vanishes, but there is no cancellation in gravitation since all the forces are of the same sign. So with the Earth made up of over 10^{50} particles, the tiny gravitational forces due to each add up to something very noticeable (we get tired when we stand for a long time!) while the vastly larger electrical forces cancel.

Yet the consequence of this smallness of gravitation is that we have no possible way of measuring how much a proton attracts another one, nor can we tell how the electrically and magnetically very active elementary particles (even a neutron has a magnetic dipole moment!) react individually to gravitational attraction by direct experiment. Even more unimaginable is a direct measurement of how the more exotic elementary particles respond to gravitation. Thus we are confined to test different ordinary materials just as Galileo, Eotvös and Dicke did. Accordingly there is a gap in our knowledge of how total the applicability of Galileo's principle is. But we certainly know nothing that contradicts it. One ingenious argument that has markedly reduced our area of uncertainty is due to Schiff who pointed out that in different materials there are different quantities of the mildly exotic particles 'virtual' positrons. This interpretation of the experiments at least shows that positrons are no different from other matter. All in all, the sensible assumption is that Galileo's principle is completely valid.

The observable of gravitation

When one falls freely all one's goods fall with one. But free fall on the Earth is very limited. However, an orbiting satellite is in permanent free fall. As is well known, in such a body the fact that everything falls together produces a state of weightlessness. The astronaut floats around the cabin, and the drops of soup float equally. Without looking out of the vehicle, the astronaut has, it seems, no inkling that gravitation acts on the spacecraft. Gravitation thus appears to have been abolished for this space traveller. However, this is not quite true. Whereas Galileo's principle states that all bodies *at the same place* fall equally fast, irrespective of their composition, it in no way asserts that bodies in different places fall equally fast. The gravitational acceleration due to the Earth diminishes with distance from it, and the spacecraft is of finite size. Thus the part of the spacecraft nearest to the Earth will want to fall a little faster than the bulk of it, the part farthest from the Earth a little more slowly. The strength of the spacecraft means that in fact it falls with a compromise acceleration as a whole, but, say, grains of dust in its air, if near the area closest to the Earth will fall a little faster and so settle in the areas nearest to the Earth, grains close to the part farthest from the Earth will fall a little more slowly and so will settle as far from the Earth as possible. Therefore by observing that the dust accumulates in the two diametrically opposed areas of the spacecraft, the astronaut can learn from *internal* evidence that the vehicle moves in a gravitational field. This is how gravitation manifests itself, as the *nonuniformity of acceleration*, as the relative acceleration of neighbouring particles. Of course we are familiar with this phenomenon. In the example given, think of the spacecraft as our Earth, of the Earth as the Sun, of the dust as the ocean, and what we observe are the solar tides (the slightly larger lunar tides originate in just the same way, but the relative sizes of bodies are very different). Thus the solar tides manifest to us that the Earth is falling towards the Sun in its orbit and that the fall takes place with a compromise acceleration which is too low for the parts nearest the Sun and too high for the parts farthest from the Sun. (In relation to the Moon, this motion about the common centre of mass, which in the solar case is inside the Sun, is in the lunar case within the Earth. This gives a slightly different picture.)

In a shore location even under the thickest imaginable cloud, the local physicists could infer the existence of the Sun and the Moon from the observation of the tides. The tidal force is the true universal observable of gravitation. The variation of gravitation with position is always there, always measurable.

For readers mathematically inclined, this characteristic can be expressed in the following form (those not so inclined should omit the next paragraph):

Since a small difference in position characterized by a small displacement vector δx^j ($j = 1,2,3$) will lead to a small difference in the acceleration vector δf^i, the relation between them must be linear and mediated by an entity called a tensor:

$$\delta f^i = a^i_j \delta x^j \text{ (where the repeated index } j \text{ is summed).} \quad (1)$$

The nine quantities $a^i{}_j$ thus define the gravitational field. To avoid the field spinning up a sphere indefinitely (which is not observed and would be an impermissible inexhaustible source of energy) the $a^i{}_j$ must have a certain symmetry. In Cartesian coordinates

$$a^i{}_j = a_{ij} = a_{ji} = a^j{}_i. \quad (2)$$

Newton's theory of gravitation is essentially the statement that the a_{ij} are minus the second derivatives of a scalar, the gravitational potential, so that

$$a_{ij} = -\frac{\delta^2 V}{\delta x^i \delta x^j}. \quad (3)$$

Poisson's equation defines the sources of gravitational field as the density of ρ multiplied by the gravitational constant G through

$$a^i{}_1 = -\nabla^2 V = -4\pi G\rho. \quad (4)$$

This completes the description of Newton's theory.

Leaving the mathematics for the time being, the question of 'negative mass' arises. Why is gravitation always attractive? First we have to define what one means by mass. The modern physicist, through bitter experience, has learned that clear meaning can be attached to terms only by defining them through the method by which they are measured. It turns out that there are three distinct ways in which 'mass' is measured and accordingly three distinct kinds of mass:

Inertial mass. This is the measure of the resistance of a body to being accelerated. For example if a hammer blow of given strength is made to hit a succession of balls of different sizes and compositions, their resulting velocities will be in inverse proportion to their inertial masses m_r.

Passive gravitational mass. This measures a body's response to gravitational attraction and is indeed its weight, as measured by, say, a spring balance. It will be called m_p.

Active gravitational mass. This measures the ability of a body to generate a gravitational effect and is ascertained thereby. Since only very large bodies produce noticeable gravitational forces, this method is largely confined to astronomy. We know the masses of the Earth, the Sun, the Moon, etc. by measuring the orbits of other bodies in their neighbourhoods. This mass is called m_a.

What can we say about the relation between these different measures of mass? Do they measure the same quantity or not? First we get guidance from Galileo's principle. Since the acceleration of free fall of a body is proportional to the gravitational force on it, which is itself proportional to m_p, and inversely proportional to its inertia and thus to m_i, it follows that m_p/m_i is the same for all bodies and, with suitable units, may be put equal to one. Next, Newton's third law of motion, of the equality of action and reaction, implied that $m_a = m_p$. But none of these relations, important as they are, shows that all mass must be positive. How would negative mass reveal itself?

Think first of a body with negative m_a and therefore negative m_i and m_p. It would repel *all* bodies whatever the sign of their mass, since by Galileo all bodies fall equally fast. If near it there were a body of positive mass of the same magnitude, then this would similarly attract *all* bodies, including the negative one. Under suitable circumstances they would both move off in the same direction, the positive one leading, the negative one chasing it, but keeping their distance from each other constant. Thus the forces would remain the same and the pair would rush off at constant acceleration. 'How absurd,' the reader may think, 'surely there must be something in physics to forbid this happening, such as the laws of conservation of energy or momentum.' Unfortunately our sanity cannot be saved in that manner. By what has been said, the body of negative m_a has also negative inertial mass m_i. Thus, since the two bodies always have equal velocity, the momentum and the kinetic energy of the negative body will always be equal and opposite to those of the positive body, so that both energy and momentum of the pair taken together always vanish. Thus they are perfectly conserved. Nor could *I* gain any energy by catching the pair since *I* would lose as much from catching the negative one as *I* gained from the positive one. Is there a way of catching the positive one without catching the negative one? Perhaps, but nobody seems to have thought through this 'thought experiment', which is very difficult.

Since we never see such a pair, it seems reasonable to conclude that no negative masses exist in our neighbourhood and indeed that it would be odd if anywhere positive and negative masses coexisted. On the other hand a universe of purely negative masses has been imagined and studied by Bonnor. It would be very different from the one we know, but it would contain no obvious absurdities. However, it is reasonable to confine our attention to positive masses.

Newton's theory has another interesting and relevant aspect which can be explored with the use of a little elementary mathematics.

Imagine a spherical body of mass M and radius R, and think of a small body in close circular orbit about it. How long will this satellite take to complete a single orbit? Calling its period P, the velocity v of the body will be the ratio of the circumference of the body, $2\pi R$, to P. In its circular orbit the satellite will have a centripetal acceleration v^2/R which must equal the gravitational acceleration GM/R^2, G being the constant of gravitation. Thus

$$\frac{GM}{R^2} = \frac{v^2}{R} = \frac{(2\pi R)^2}{P^2 R} = \frac{4\pi^2 R}{P^2}. \quad (5)$$

Accordingly

$$P = 2\pi (R^3/GM)^{\frac{1}{2}} = 2\pi (3/4\pi G\rho)^{\frac{1}{2}}, \quad (5')$$

where ρ is the mean density of the body, i.e. the ratio of its mass M to its volume $4\pi R^3/3$.

Thus, after deriving Kepler's third law for circular orbits (it holds equally for elliptical ones, with the semi-major axis taking the place of the radius R), this equation shows that the period of a close satellite depends only on the *density* of the parent body.

Imagine now a series of such spherical bodies all having the same mean density but each being larger than its predecessor. The period of their close satellites are all the same, their mean densities being the same. But the larger the body, the faster the satellite has to go to encircle its parent body in the same interval of time. Whatever density is chosen there will be a size of body at which the satellite has to move with the speed of light! (Of course this is an extrapolation of Newtonian theory, the validity of which is restricted to slowly moving bodies, but this extrapolation is none the less illuminating.)

Indeed, since the speed of a particle projected from the surface needs to be $\sqrt{2}$ times the circular velocity, even for a slightly smaller body of this density, particles moving at the speed of light (and, by inference, light) cannot get away from such a body and so we have constructed a 'black hole'! More precisely we have arrived at the Newtonian analogue of a black hole.

Note that for any mean density there will be a radius (and accordingly a mass) for which the body will be a black hole. Thus a black hole need not have a high density, but the lower its density, the bigger the mass has to be. It is easily seen from the equation given that the minimum mass necessary is proportional to $\rho^{-1/2}$. A body of the mean density of our Sun would need to have 500 times its radius to become a black hole!

Relativity and gravitation

Newton's dynamics and theory of gravitation are splendid pieces of physics, but tell us nothing about light. This is a very serious defect since light (and other electromagnetic radiation) constitute our chief means of carrying out observations. In the absence of gravitation, light and all radiation, together with its interaction with matter, is described by Einstein's special theory of relativity (the term 'special' means just that it does not describe the situations in which gravitation is in any way relevant). For the purposes of this article it will be assumed that the reader has a nodding acquaintance with the theory of relativity. In particular, it will be necessary for the reader to recall two important conclusions of it, both amply confirmed by countless experiments: (i) energy has mass, i.e. it has inertia, (ii) the law of composition of velocities ensures that it is impossible to accelerate a body (moving necessarily with a speed less than that of light) to the speed of light or beyond.

For the mathematically inclined, there is a third result that will be useful: If two events are observed by one inertial observer to take place in positions with coordinates (x_1, x_2, x_3), $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ respectively at times t, $\bar{t}$ respectively then a differently moving inertial observer will assign different space and time coordinates to them, but the quantity

$$(t-\bar{t})^2-[(x_1-\bar{x}_1)^2+(x_2-\bar{x}_2)^2+(x_3-\bar{x}_3)^2]/c^2$$

will be found to be the same by both inertial observers (here c is the speed of light). This statement is analogous to the statement that in a plane the coordinates of two points will have different values in two differently oriented systems of coordinates, but the distance between the points can be evaluated from both giving the same result.

Returning now to the nonmathematical description, we owe Einstein a most illuminating thought experiment on the interaction between gravitation and light. First the reader should be reminded of an important and well-established result of atomic physics. An atom of any element has a ground state and a number of excited states. If it is in one of these excited states, it can make a transition to the ground state by emitting light of a particular frequency, i.e. a photon of a particular energy. If light of this particular frequency is absorbed by the atom in the ground state it makes the transition to the particular excited state mentioned before. It will be evident to the reader that the atom in the excited state has more energy than when it is in the ground state, since it can emit a photon of particular frequency.

Now consider the following situation: A closed chain supporting numerous buckets is stretched between a wheel at the top of a tower and one at its foot. Each bucket is filled with the same number of atoms of the same element, but all the atoms on the left side are in the excited state mentioned while all the atoms on the right hand side are in the ground state. Thus all the atoms on the left hand side have more energy than those on the right hand side, and accordingly more mass and therefore, by Galileo's principle, more weight. Hence the left hand side is heavier and will move downwards, while the right hand side moves upwards. The excited atoms arriving at the bottom are induced there to make the transition to the ground state, emitting light as described. A set of fixed mirrors is so arranged that all the light emitted at the foot of the tower is directed to its top, there irradiating the atoms in the ground state arriving as the right side of the chain moves up. Then, by what has been said, the light shining on them would put the atoms from the ground state into the excited state before they start on their downward journey on the left hand side. Thus, it seems, the system would always be in the same condition, the left hand side heavy and going down, while the right hand side of the chain is lighter and goes up. Thus the chain goes round and round, and could be used to drive a generator!

Plainly, this cannot be, for such a system would generate energy out of nothing, contrary to the well-established law of conservation of energy! There must be a flaw in the argument given, but where can it be? Given that we know from both experiments and theory that there are excited states, that more energy means more mass, and that more mass means more weight where can the fault in the argument lie? The only possibility is that while the light emitted by the transition to the ground state of the excited atoms has just the right energy (i.e. frequency) to excite atoms in the ground state *at the same height*, this is no longer so when the atoms to be excited are higher up. Then, we assume, the light is too low in energy, i.e. too low in frequency, to excite atoms there. Thus we arrive at the idea of a *gravitational red shift* of light. With this assumption the light getting from the bottom to the top would be unable to excite the atoms in the ground state arriving there, and so the paradox would not arise.

We can readily complete our thought experiment to show that this explanation adequately resolves the energy problem in our model. For if light is reflected from a mirror advancing into the beam, it is blue-shifted. So we arrange to have a spinning wheel of mirrors at the top of the tower. If it spins at the correct speed, this reflection from moving mirrors compensates for the gravitational red shift, and the light is again enabled to excite the ground state atoms arriving at the top. However, light exerts a pressure, and so it consumes energy to turn the wheel of mirrors, just as much energy as the chain produces. Thus the inference of a gravitational red shift resolves our problem: The model neither generates nor consumes energy, in full agreement with the law of conservation of energy.

As soon as Einstein stated his deduction of the necessity of a gravitational red shift there was great interest in testing his idea by experiment and observation. However, the shift is very small. The largest one in our neighbourhood occurs for light from the Sun, the spectral lines of which are generated in the Sun's atmosphere. Light from the Sun 'ascends' from there (i.e. it moves against the Sun's gravitational force) until it reaches the neighbourhood of the Earth to whose surface it finally descends. (However, the blue shift it suffers in this descent is well below one thousandth of the red shift acquired in the ascent.) Thus in a terrestrial laboratory the spectral lines in the light of the Sun should, according to Einstein, be slightly shifted towards the red compared with the same lines produced in the laboratory, and this shift should be just about observable. However, in spite of much painstaking and skilful work (especially by M. G. Adams in Oxford during the fifties), the results were not wholly decisive. The trouble arose through quite secondary shifts of spectral lines caused by conditions in the Sun's atmosphere that could not be reproduced in the laboratory and could not, at the time, be computed. It will also be evident that the precision with which the position of a spectral line can be measured must depend on how wide or diffuse the line is. In the late fifties, in a wholly separate field of physics, Mössbauer discovered a way of producing exceedingly narrow gamma-ray lines, far narrower than had previously been considered possible. In 1960 Pound and Rebka of Harvard used this Mössbauer effect to compare at the top of a water tower a Mössbauer line generated there with one produced at the foot of the tower, and indeed measured the Einstein red shift exactly as predicted, though of course it is quite minute in such circumstances (one part in 300 million million). This experiment (repeated since with ever greater accuracy) fully establishes the reality of the gravitational red shift.

Its importance is not diminished by its smallness. Spectral lines are the basis of time keeping. The most accurate clocks are in the standard laboratories round the world, using exceptionally stable spectral lines of caesium or ammonia. The most precise watches of our day use a spectral line (in the MHz range) of a quartz crystal. Earlier watches used the oscillatory frequency (i.e. spectral line) of a spring. Pendulum clocks are excluded because they are not of universal applicability, requiring as they do a gravitational field. Even on the surface of the Earth their operation varies with altitude and location.

Thus what the gravitational red shift means is quite striking: Take two watches of identical construction. They will keep the same time if kept side by side. But if one is in a higher position than the other, they will no longer keep in step! Though the difference in tick rate may be small, it is of fundamental significance.

Physicists have learned from much experience not to ascribe any absolute existence to any of their concepts. Just as it was necessary to define mass by the method by which it is measured, so we cannot assert the existence of the 'right' time in any sense. Time is that which is measured by clocks (or we can say 'manufactured by clocks') and if identically constructed clocks have different tick rates in different positions, then time is different in the two locations. There is no sense in which one can assert that the higher positioned clock gives a 'right time' and the lower positioned one a 'distorted time'. One is as good in one location as the other is in the other one, but time itself is different in the two places.

It is instructive to look at this graphically. Draw time, as is conventional in physics, vertically up and the one relevant space dimension, height, horizontally. One vertical line then represents the foot of the tower, another the top, and suitably sloping parallel lines the travel of light signals from the bottom to the top. Consider now the depiction of two successive such signals. The diagram shows a parallelogram, but the time interval between the signals, measured at the foot by the local clock there and correspondingly at the top, will have different magnitudes at the bottom and at the top because of the different rates of the two clocks. Thus our parallelogram has unequal opposite sides. This is incompatible with Euclidean geometry. Thus a non-Euclidean geometry must hold in the two-dimensional space of time and height. Two-dimensional non-Euclidean geometry need not frighten one, since we are used to it from the curved surface of the Earth. However, if we add to height the two horizontal dimensions of a (nearly) spherically symmetric Earth, e.g. latitude and longitude, then it turns out that a height-dependent time is incompatible with a Euclidean space-time.

In mathematical terms we are dealing with a spherically symmetric space-time

$$ds^2 = f(r)\,dt^2 - g(r)dr^2 - r^2h(r)(d\theta^2 + \sin^2 d\phi^2), \qquad (6)$$

in an obvious notation.

It is readily established that such a space-time cannot be flat (Euclidean) unless $f(r)$ is a constant. Since there is a gravitational red shift, $f(r)$ varies with height r, and so space-time is non-Euclidean.

It is worth pausing for a moment to look back at the chain of argumentation employed. Starting from Galileo's principle, we added the framework of special relativity (energy has mass, time and space) and the existence of spectral lines. This is enough to infer that gravitation leads to a non-Euclidean nature of the world's geometry.

The mathematical theory of gravitation (general relativity)

Return to the earlier mathematical formulation of Newton's theory, viz. equation 1

$$\delta f^i = a^i_j \delta x^j \text{ (where the repeated index } j \text{ is summed).} \quad (1)$$

This is clearly incompatible with special relativity, since this describes a velocity-independent acceleration. If the reference particle were moving at nearly the speed of light, another particle displaced from it in a suitable direction could be accelerated through the speed of light, a process incompatible with our knowledge of physics. Thus a'_j in equation 1 must be velocity dependent.

The easiest way to achieve this is to regard equation 1 as a four-dimensional equation and put the velocity v^k into the picture through

$$a'_j = b'_{jk}v^k. \quad (7)$$

With b antisymmetric between i and k, this would, through the algebra of special relativity, avoid any possible transgression of the speed of light. Unhappily antisymmetry between i and k is incompatible with symmetry (of equation 2) between i and j. We are thus driven to the next more complicated form,

$$a^i_j = C^i_{jkl}v^k v^l. \quad (8)$$

With symmetry between i,j and between k,l, and antisymmetry between i and k (and between j and l) all the conditions are met. Transgression of the speed of light is impossible, and for low velocity (for which only the time component of v differs significantly from zero) the Newtonian equation 1 is regained. Thus equation 8 extends equation 1 naturally and easily from the low velocities well described by it to any velocity up to (and including) the speed of light. This major progress is achieved at a cost. Instead of six free components of the symmetrical three-dimensional two-suffix tensor a we have the four-dimensional four-suffix tensor C. Its complicated symmetries ensure that of the basically 256 components of such an entity only 21 can be chosen freely. But it is not surprising that the huge increase in the complexity of the picture described involves a growth in the number of free components from six to 21.

Now return to the non-Euclidean character of space-time demonstrated by equation 5. The least complicated non-Euclidean geometry is Riemannian. In this geometry a most important entity is one describing the deviation of space-time from flatness. This is the curvature tensor

$$R^i_{jkl}. \quad (9)$$

We need not discuss here how it is evaluated, but point out its essential symmetries which (but for a different convention in the arrangement of the suffixes) are exactly as for the C^i_{jkl} of equation 8 but with one additional identity (reducing the number of free components to 20) and a set of differential identities governing the variation of R from point to point. Thus the vital equation 8 describing the relative acceleration of neighbouring particles, the fundamental observable of gravitation, finds a beautiful geometrical interpretation by identifying C with R. This effectively replaces equation 3. But what about equation 4 which links gravitational field to its sources? How can its relativistic equivalent be found?

First, it is plausible that in a relativistic theory the active gravitational source should depend not only on matter itself but on energy, momentum and stress. A box full of radiation should have a different gravitational effect than an empty box. Moreover whereas in Newtonian theory the permanence of its source, mass, was an extraneous addition to the theory, it would be agreeable if the conservation characteristics of mass and momentum could be linked to its gravitation-causing character. Einstein achieved all this brilliantly.

First, he described the source of gravitation through the so called energy-momentum tensor T^{ij}. It is most easily visualized in a tenuous cold gas where

$$T^{ij} = \rho v^i v^j. \quad (10)$$

ρ being the density of matter and v^i, as usual, its velocity vector. It will be recalled that in Cartesian coordinates, its time component (according to special relativity is

$$v^0 = [1 + (v^1)^2 + (v^2)^2 + (v^3)^2]^{1/2}. \quad (11)$$

Thus T^{00} represents the mass density enhanced by the kinetic energy, T^{01} the 1 component of momentum, while the other components also have significance. Proceeding to a dense hot gas, mean values over the velocities have to be taken so that the space–space components represent stresses and pressures. This is then readily extended to elastic solids.

Next Einstein constructed a symmetrical two-suffix tensor G^{ij} whose components are a linear combination of components of the curvature tensor R and finally he put

$$G^{ij} = -kT^{ij}, \quad (12)$$

where k is a constant depending on the choice of units. This corresponds closely to equation 4, where in the Newtonian scheme a linear combination of the a^i_j was put equal to the source, the density of matter. But quite apart from the ability to describe fast moving matter and light, equation 12 has another superiority over equation 4. For the Einstein tensor G^{ij}, through the differential relations of the curvature tensor, satisfies a kind of conservation law which by equation 12 implies similar conservation for the energy momentum tensor T^{ij}. This completes our mathematical description.

First results of the theory

Our solar system is very amenable to calculation. The masses of the planets and their satellites are so small compared to the mass of the Sun that one can first work out the orbit of each planet as though the Sun was the only attracting body, and then add the influences of the other bodies as minor perturbations. It is perhaps not surprising that for slowly moving bodies, as most planets are, the relativistic orbits are virtually identical with the Newtonian ones, but of course relativity allows us to treat light and fast moving bodies as well. The only fast moving planet is Mercury, and relativity readily leads to a deviation from the Newtonian orbit that had been known and was a puzzle for many decades before. Einstein's theory of gravitation fully accounts for this orbital feature. The Einstein gravitational red shift has already been discussed. The effect of the Sun on a passing ray of light can only be described relativistically. Although the forecast of the theory was for long difficult to test astronomically, it has been strikingly verified by the observation of space probes that go behind the Sun.

It is most gratifying that the theory has passed all these tests so well, for general relativity is the simplest acceptable theory of gravitation. These tests were checks on general relativity, and did not offer a choice of going back to Newtonian theory since this, with its absolute time, is not just unacceptable but indeed unimaginable to a modern physicist. One can use Newtonian theory not as a description of nature, but as a tolerable approximation when velocities are low, gravitational fields are not too strong, etc.

The Newtonian approach to black holes was described earlier. In general relativity it is a somewhat easier concept. If we look from a distance at a sphere of given density and large radius, there will of course, be a sizeable red shift to be observed. The larger the sphere, the bigger the red shift, i.e. the slower things seem to be happening there, the less the amount of light received. For a certain radius, the red shift will be total. We will therefore receive no light from such a 'black hole'.

Gravitational energy

Energy is perhaps the most fundamental and universal concept in physics. Energy occurs in many different forms: The energy of motion (kinetic energy), heat, strain energy of an elastic body, pressure energy (e.g. of steam), the chemical energy of a combustible substance, electric energy, the latent energy of an excited atom (as previously discussed), etc. etc. However, there are some difficulties with the concept of gravitational energy. Even in the Newtonian theory there are some problems, as will now be described.

The simplest case perhaps is that of the weight in a pendulum clock. When the weight is in a low position, energy has to be expended to raise it to a high position. Once it is there, it can descend slowly, giving energy to the clock mechanism to make it work. Thus energy has gone from the winder to the weight and then into the mechanism. The weight has stored this energy not by changing its composition, not by being hotter, but by being *higher up* in the gravitational field of the Earth. Thus this is an energy of position or, as it is generally called, a *potential* energy. (It is closely related to the V of equation 3.) Though strictly speaking this is an energy of interaction between the weight and the Earth, there is little harm in ascribing it essentially to the weight and supposing it to reside in the weight, though there is nothing to check this assumption.

Next let us consider a somewhat more complex system, a pair of stars in elliptical motion about their common centre of mass. At one moment they are far from each other, moving relatively slowly, then their distance diminishes while their speeds increase in this descent towards each other, and their velocities reach a maximum at closest approach, diminishing thereafter as their distance increases. Thus their kinetic energies are at a minimum when they are farthest from each other and at a maximum when they are nearest to each other. By conservation of energy, there must be a potential energy going through the opposite variations so that the sum of the kinetic and potential energies remains constant. This potential energy is evidently a property of the *system* and cannot be assigned meaningfully to one or the other body in any sensible proportions. Thus this potential energy has no specific location. This is perfectly acceptable in Newtonian theory where potential energy is merely a useful mathematical construction to make the books balance. But it will be readily understood that in relativity in which energy inevitably has mass, a nonlocalizable form of energy is wholly unacceptable.

Before we leave the discussion of energy in the Newtonian theory, some further points can be made especially about the transfer of energy by gravitation.

In the early part of this article it was shown how the Moon raises tides on the Earth. Since the Earth rotates, these tidal bulges (of the ocean, the solid Earth and the atmosphere) will rub against the Earth just like a brake shoe rubs against a wheel. This tidal friction does indeed slow down the Earth's rotation. A few billion years ago, the Earth probably had a rotation period of only about 8 hours, spinning three times as fast as today. With our very precise clocks, we can notice the slowing down in our time. Moreover, this tidal friction affects the motion of the Moon. The tidal bulges, being dragged forward by this friction, themselves exert a gravitational force on the Moon dragging it forward in its orbit, and thus very gradually driving the Moon further away from the Earth (though the Moon must at some stage in the past have been much closer to the Earth, it is reasonably clear that the Moon did not originate from the Earth). Alternatively it is evident that the angular momentum of the Earth Moon system cannot change. The diminution of the Earth's spin must be compensated by an increase in the Moon's orbital angular momentum, equally making clear its gradual recession from the Earth.

All this can be checked from the historical record of eclipses of the Sun and the Moon. Especially a solar eclipse through its location on the surface of the Earth is most informative about changes in the Earth's period of rotation, in the Moon's orbit, and even in the moment of inertia of the Earth. Initially the Earth raised large tides of the solid material of the Moon. The internal friction of this squeezing of its body has been so large that the Moon's spin has been slowed down to the period of its revolution about the Earth. This is why the Moon always presents the same face to the Earth.

Evidently there are intricate transfers of energy within the Earth–Moon system, but it is amusing to construct mentally a kind of toy that demonstrates such gravitational energy transfers with total clarity. Let us again imagine two originally spherical bodies A and B in highly eccentric orbits about their common centre of mass, and let us further imagine that on each of these bodies there are engineers capable of changing the shapes of the bodies at will, but keeping them to be axially symmetric about the normal to the common plane of the orbits. Thus either body can be changed from being a sphere to being an oblate or prolate spheroid. Each body, moreover, has electric storage batteries so arranged that by reducing their charges they can energize the shape changing machinery where such a change requires energy, and increase their charges where the change of shape yields energy.

Three points need to be made now: An oblate body attracts more strongly in its equatorial plane than a sphere of the same mass, a prolate body less so. Secondly the tide raising force, which aims to lengthen a body along the line joining the two, favours an oblate over a prolate shape. Thirdly the tide raising force diminishes greatly with distance.

Suppose now that the engineers on A persuade their colleagues on B to make always the opposite changes to their own object. Thus whenever A goes prolate, B goes oblate and vice versa, in such a manner that the mutual gravitational attraction between the two bodies always remains the same, as when they both are spheres. Thus the orbits will remain always the same. Now suppose that in one revolution A changes its shape from a sphere to an oblate spheroid when the two bodies are in proximity. There the tide raising force is great, and favours this change of shape so that the operation yields a substantial amount of energy greatly increasing the change on A's battery. When the bodies are at their maximum distance apart. A returns to the spherical shape. This will cost it energy, but not very much because when B is so far from A, the tide raising force is very weak. So although the second change of shape reduces the charge on A's batteries, it will be much less than the gain experienced earlier. Thus at the end of the revolution, A has gained energy. Naturally B's experience is the opposite. He had to go prolate when this was most expensive as in proximity A's tide raising force on him is so strong. His battery will be badly discharged in this manoeuvre. Of course he gains some energy when he changes from prolate to spherical at maximum separation but not very much because the tide raising force is then so weak. Thus at the end of the revolution, the situation is just the same as at the beginning, except that A's batteries carry more charge than before, and B's less. Thus energy has been transferred from B to A purely by gravitational means.

Now let us leave Newtonian considerations, however elegant, behind us, and see what reality one can give to the energy concept in relativity. As has already been stated, potential energy can have no place in a relativistic theory of gravitation because it is nonlocalizable. Since mass is a necessary consequence of energy, and a nonlocalizable mass is an inadmissible absurdity, we have to manage without potential energy, although this was needed in Newtonian theory in order to get an energy balance. The energy–momentum tensor introduced earlier only encompasses nongravitational forms of the energy which may be called tangible. Since tangible energies are not conserved (remember the weight in the pendulum clock), Einstein's law of conservation is really a law of non-conservation. But not only is the conservation of energy a cornerstone of physics but we know that the Newtonian equations are an acceptable approximation to the relativistic (and far more complicated) equations for slow motions, modest gravitational fields and pressures small compared to densities. Thus somehow in relativity conservation must come in through the back door. The truth of the matter is that although general relativity has been our theory of gravitation for three quarters of a century, its concepts and equations are so different that there is still much we do not understand in any detail.

Though the relativistic model of A and B is beyond our mathematical competence one can investigate a similar model which has axial symmetry, A being a spheroidal shell surrounding the spheroidal solid B. One can then indeed demonstrate that slow changes of shape shift energy between A and B, but the sum total as expressed by the external gravitational field, is strictly conserved. So conservation applies but only to systems as a whole as long as motions are slow. Much still remains to be clarified in the field, but the overall principles seem to be intelligible.

Gravitational waves

An innocent question 'How fast does gravitation travel?' can lead us into illuminating though only partially solved problems that concern fast motions, a topic so far carefully avoided. The physicist will regard this question as meaningless until a sensible model has been specified in which conceptually consistent (but not necessarily practicable) observations would indicate the answer. Perhaps one's first reaction to this attitude would be: 'Suppose the Sun suddenly ceased to exist, how soon would the Earth leave its orbit? Some $8\frac{1}{2}$ minutes later when we would see the disappearance of the Sun, or earlier, or later?' This is still a meaningless question since general relativity, our theory of gravitation, incorporates the law of conservation of mass, and so must rule out the sudden abolition of the Sun. The next more refined question might be: 'Suppose the Sun were suddenly to acquire a velocity at right angles to the plane of Earth's orbit, how soon would the Earth leave its orbit?' Again this question must be ruled out of order since our theory encompasses the law of conservation of momentum. But at the next level of sophistication the theory is obliged to give an answer: 'Suppose the Sun rapidly changed its shape from spherical to spheroidal. How soon would the Earth's orbit be affected?' Indeed we know how to answer this question. First, it will be recalled, when the Sun changes to become an oblate or a prolate spheroid, the attraction in its equatorial plane would be affected. Secondly, there would be no inkling of this change gravitationally before light from the new shape reaches the Earth, and then, some $8\frac{1}{2}$ minutes after the change took place, the Earth's motion would begin to be affected. Though most of the change would occur then, there would be some aftermath too, a kind of gradually diminishing reverberation. Thirdly, if after this excursion into a spheroidal shape the Sun then resumed its normal spherical form, there would be a certain change from what things were like before these events.

This last point is not quite simple. As the Sun changes its shape, its gravitational influence alters, its variations propagating with the speed of light (with minor squiggles following rather more slowly). The mathematical physicist calls such propagating fluctuations *waves*, whether or not they are of sinusoidal shape. These gravitational waves thus *signal* information about the Sun's shape to all and sundry. It is a general rule in physics that information cannot be transmitted without some energy. Therefore the Sun, in changing from spherical to spheroidal shape and back, must have lost energy and therefore mass. Whatever the shape of the fluctuations, the resulting radiation must have reduced the mass of the Sun, and in the final situation it must be a lesser source of gravitation than it originally was.

To put it mathematically, the measure of the spheroidal deviation from spherical shape is called the quadrupole moment Q. Then the effect on the Sun's mass M turns out to be given by

$$\frac{dM}{dt} = -\lambda\left[\frac{d^3Q}{dt^3}\right]^2 \tag{13}$$

where λ is a universal constant that, if suitable units are chosen, is purely numerical. Several points emerge from this equation. First it makes patent that as has been said, radiation of gravitational waves inevitably leads to a mass loss, and not to a mass gain. Secondly it demonstrates that our theory of gravitation necessarily is nonlinear, which is the prime cause of making the evaluation of its mathematical consequences so complex, that much still remains to be explored. Third it shows that the rate of mass loss or, to put it more positively, the amount of energy carried away by the gravitational waves depends on a high time derivative. If the fluctuations are sinusoidal the rate of mass loss goes with the *sixth* power of the frequency.

While all this is theoretically fairly complete and fully self-consistent, it leaves wide open the question of what strength of gravitational waves are travelling round the universe in fact. Fortunately for us, the Sun shows no tendency to change its shape, but a distant observer would notice changes in the quadrupole moment of the solar system due to the revolution of planets round the Sun. Jupiter, by far the most massive of the planets, would dominate this, but the period of its orbit is so long ($10\frac{1}{2}$ years) and therefore the frequency of the variation of quadrupole moments so low that this whole huge solar system turns out to radiate a mere $1\frac{1}{2}$ kW in gravitational waves! If there are powerful sources anywhere in the universe they must involve huge masses in rapid motion. One's thoughts turn to double stars. One formed of two ordinary stars like our Sun almost in contact would have a period of revolution of a few hours. Yet even so, radiation by gravitational waves would have only a tiny fraction of the energy radiated by these stars as light and heat. Shorter periods are needed which means the stars have to be closer and therefore far denser than our Sun is. The densest stars we can conceive of are composed of neutrons, and some such neutron stars manifest themselves to radio astronomers as pulsars. Two such neutron stars in close proximity forming a double star would be tremendous radiators of energy (and angular momentum) by gravitational waves. Since the mid 1970's we have known a pulsar whose signals can best be interpreted as originating from a double star system of very high frequency. Moreover, minute changes in characteristics revealed by radio astronomy are exactly those that our theory of gravitation would predict for such a system. Thus we have an observational test of the theory that confirms its predictions for the emission of gravitational waves, one of its most sophisticated consequences.

What about the reception of gravitational waves? Variations in stresses and distances could reveal their incidence on Earth. Undoubtedly their strength is minute, but a great deal of thought is going into the design of exceedingly sensitive receivers of gravitational waves. Once they are built, and make detections, a new window will have been observed on the universe: The era of gravitational wave astronomy will have begun.

Conclusion

This brief description of the field of gravitation has involved an introduction to our modern theory of gravitation, Einstein's general theory of relativity. The approach used has been largely non-mathematical but I hope has given the reader an impression of this area of science in which so much still remains to be explored.

Food prospects in India by the turn of the century

L. V. Venkataraman

The continuing population pressure demands greater efficiency in food production systems. Self efficiency in food front by 2000 AD can be achieved only by specific shifts in farm strategies. Coarse grains, the mainstay of rural population, needs greater attention. The losses of grains in field and storage are to be minimized. Biotechnological methods to augment food supplies should complement traditional technologies.

Today, India is the fastest growing country in the world. Our population is projected to grow even higher than that of China and is likely to touch the billion mark by the turn of the century. This alarming growth in population will in turn increase the demand for food. Meeting this need will require increase in food production, decrease in food losses, improvements in food processing, and enhanced nutritional quality and safety of processed foods.

World Development Report (WDR) 1990 has stated that there is considerable potential for progress towards reducing poverty in India. For this, higher investment with better domestic saving rates and external borrowing will be needed. India's growth rate is unlikely to exceed 2% per capita a year over the decade with 370 million people below the poverty line by 2000 AD. Between 1985 and 2000, the incidence of poverty in the developing world has been projected to fall from 33 percent to 18 percent and the number of poor from 1.1 billion to 82.5 million (Table 1).

The food problem cannot be separated from the challenges of population growth and efficient organization of food production systems. Current trends in population growth, food production and food consumption cannot continue; they need to be looked at very carefully. Population growth must be controlled if massive famines are to be avoided in a predominantly monsoon-dependent country like India.

The technological developments that are likely to have the greatest impact on food production should also emphasize simplicity, low cost, labour-intensiveness and appropriate technology. The fickleness of weather still determines whether the poorer section of the population will meet its food demand or face starvation. Our country will need to give maximum attention to agriculture in order to

1. assure basic food security to the people,
2. improve abysmally low nutritional standards and
3. create surpluses for poverty alleviation programmes.

Table 1. Poverty in 2000 AD

Region	Incidence of poverty (%) 1985	Incidence of poverty (%) 2000	Number of poor (million) 1985	Number of poor (million) 2000
China	20.0	2.9	210	35
India	55.0	25.4	420	255
Developing countries	32.7	18 0	1125	825

Based on world development report. Poverty line at $ 370 annual income.

L. V. Venkataraman is a senior scientist, Central Food Technological Research Institute, Mysore 570 013, India