

When Newton saw an apple fall, he found
In that slight startle from his contemplation -
'Tis said (for I'll not answer above ground
For any sage's creed or calculation) -
A mode of proving that the world turned round
In a most natural whirl, called "gravitation";
And this is the sole mortal who could grapple,
Since Adam - with a fall - or with an apple.

Men fell with apples, and with apples rose
If this be true; for we must deem the mode
In which Sir Isaac Newton could disclose
Through the then unpaved stars the turnpike road,
A thing to counterbalance human woes;
For, ever since, immortal man hath glowed
With all kinds of mechanics, and full soon
Steam-engines will conduct him to the moon.

George Gordon, Lord Byron (1788-1824)
Don Juan, Canto X

When Newton saw an apple fall, he found
 in that slight startle from his contemplation -
 'Tis said (for I'll not answer above ground
 - For any sage's creed or calculation)
 A mode of proving that the world turned round
 In a most natural whirl, called "gravitation";
 And this is the sole mortal who could grapple,
 Since Adam - with a fall - or with an apple.

Men fell with apples, and with apples rose
 It this be true; for we must deem the mode
 In which Sir Isaac Newton could disclose
 Through the then unopened stars the turnpike road,
 A thing to counterbalance human woes;
 For, ever since, immortal man hath glowed
 With all kinds of mechanics, and full soon
 Steam-engines will conduct him to the moon.

George Gordon, Lord Byron (1788-1824)
Don Juan, Canto 8

note to student:

this is conceptually very difficult
but mathematically relatively easy

Special Relativity

Besides quantum mechanics, classical mechanics received in the twentieth century another major modification. This modification was special relativity, as formulated by Albert Einstein. Its basis is just two ideas: one old, one new (Box 3.5). The old idea is the relativity principle.

elsewhere + more.

BOX 3.5

The Basic Postulates of Special Relativity

- (a) The relativity principle: the laws of physics are the same for all inertial observers.
 - (b) The constancy of the speed of light: the speed of light, $c = 3.00 \times 10^{10}$ cm/sec, is the same for all inertial observers, independent of their velocity or motion relative to the source of the light.
-

Shu, The Physical Universe

We have already discussed, in connection with electromagnetism, the concept that the laws of physics (not necessarily the *appearance* of phenomena) are the same for all inertial observers. They do not depend on the observer's velocity with respect to the "fixed stars." The new and astounding idea is that the speed of light in a vacuum, c , is the same for all inertial observers. In other words, the speed of light equals 3.00×10^{10} cm/sec independent of how fast the observer moves relative to the source of the light!

Consider what this second idea means. I shine a flashlight and observe the photons to rush away from me at 3.00×10^{10} cm/sec. You move along the path of the outwardly rushing photons at half the speed of light relative to me. I ask you how fast the photons are rushing past you, fully expecting to hear " 1.50×10^{10} cm/sec." Imagine my surprise when you tell me " 3.00×10^{10} cm/sec"! Imagine the world's surprise when Michelson and Morley told them " 3.00×10^{10} cm/sec"!

How can this answer be right? Yet, right it is, checked time and time again in many different experiments. How do we reconcile this answer with our common-sense notion of space and time? It took the genius of Einstein to see the way:

Space and time cannot be thought of as separate entities. They are intrinsically linked, and it is only a combination of the two which will appear the same to two observers who move relative to one another.

Indeed, speaking colloquially, we may say that the great achievement of Einstein's special relativity theory is precisely this unification of space and time. This unification has several well-known astounding implications, which we shall proceed to develop in a series of problems now, using no mathematics more complicated than high-school algebra and geometry. It is a fallacy that Einstein's theory of special relativity is difficult to understand. Indeed, David Hilbert, the famous mathematician, was fond of saying that every schoolboy in Göttingen could derive special relativity. And so they can—after Einstein has shown the way, and we accept the two basic postulates of special relativity (Box 3.5).

Before we begin our discussion, I should remark that the preferred modern derivation of the results of special relativity proceeds by a different route, one that makes explicit use of the four-dimensional transformation properties of spacetime. However, the concise derivations possible with such a formalism can be appreciated only after one has already developed some intuitive feeling for relativity theory, and I believe it is best developed by retracing the historical route discussed below.

Time Dilation

The first result we wish to derive is the concept of **time dilation**. This concept can be illustrated forcefully by the so-called twin paradox (see Figure 3.25):

Once there were a woman and a man who were twins. The woman left on a rocket ship for a trip to Alpha Centauri, traveling at 86.6 percent of the speed of light. The man stayed at home. The man aged five years before the woman arrived at Alpha Centauri (4.34 light-years from the solar system). The woman aged only half as much, two and a half years.

How can this be? The answer is that the man who stayed "stationary" thinks that time is "dilated" for the woman who travels relative to the "fixed" stars: all physical processes seem to slow down for her, including her biological clock. According to the woman, however, everything is perfectly natural; the pace of life and everything else proceeds on the rocket trip just as it did at home on Earth. She brought along a clock just to make sure; to make our reasoning simple, we shall suppose this clock to be constructed as follows. It is a "light" clock in which a ray of light bounces back and forth between two flat mirrors separated by a distance d . Each time the light strikes one of the mirrors, the clock makes a ticking sound; so, by multiplying the number N of ticks by the distance d and dividing by the speed of light $c = 3.00 \times 10^{10}$ cm/s, the woman can keep track of the passage of time, t_0 . The man hears (through a radio link) the same number of ticks, N , but for him the light ray must make a zigzag path in order to bounce between two *moving* mirrors. Thus, the total path length traveled by the light ray to make N ticks appears longer to the man than the path length Nd according to the woman. Since the speed of light is the *same* for both of them, the time elapsed for the man, t , seems longer than the time elapsed for the woman, t_0 . A simple geometrical construction (Problem 3.10) yields the quantitative relation:

$$t = \gamma t_0, \quad (3.11)$$

where

$$\gamma \equiv (1 - v^2/c^2)^{-1/2} > 1$$

and v is the speed of the woman. Notice, especially, that the woman thinks that the clock works normally, just like at home!

Problem 3.10. Show from Figure 3.25 and Pythagoras's theorem (Appendix B) that each diagonal path traveled by the ray of light as seen by the man is given by $(d^2 + s^2)^{1/2}$, where s is the distance traversed (perpendicular to the mirror arrangement) per tick. Clearly, if t is the total elapsed time for the man, $s = vt/N$. Set now the total path length $N(d^2 + s^2)^{1/2}$ equal to ct , and solve for t . Note that Nd/c is t_0 and derive equation (3.11). Verify that if $v/c = \sqrt{3}/2 = 86.6$ percent, t is twice t_0 . How much time dilation is there if you fly in a jet plane at 3×10^4 cm/sec?

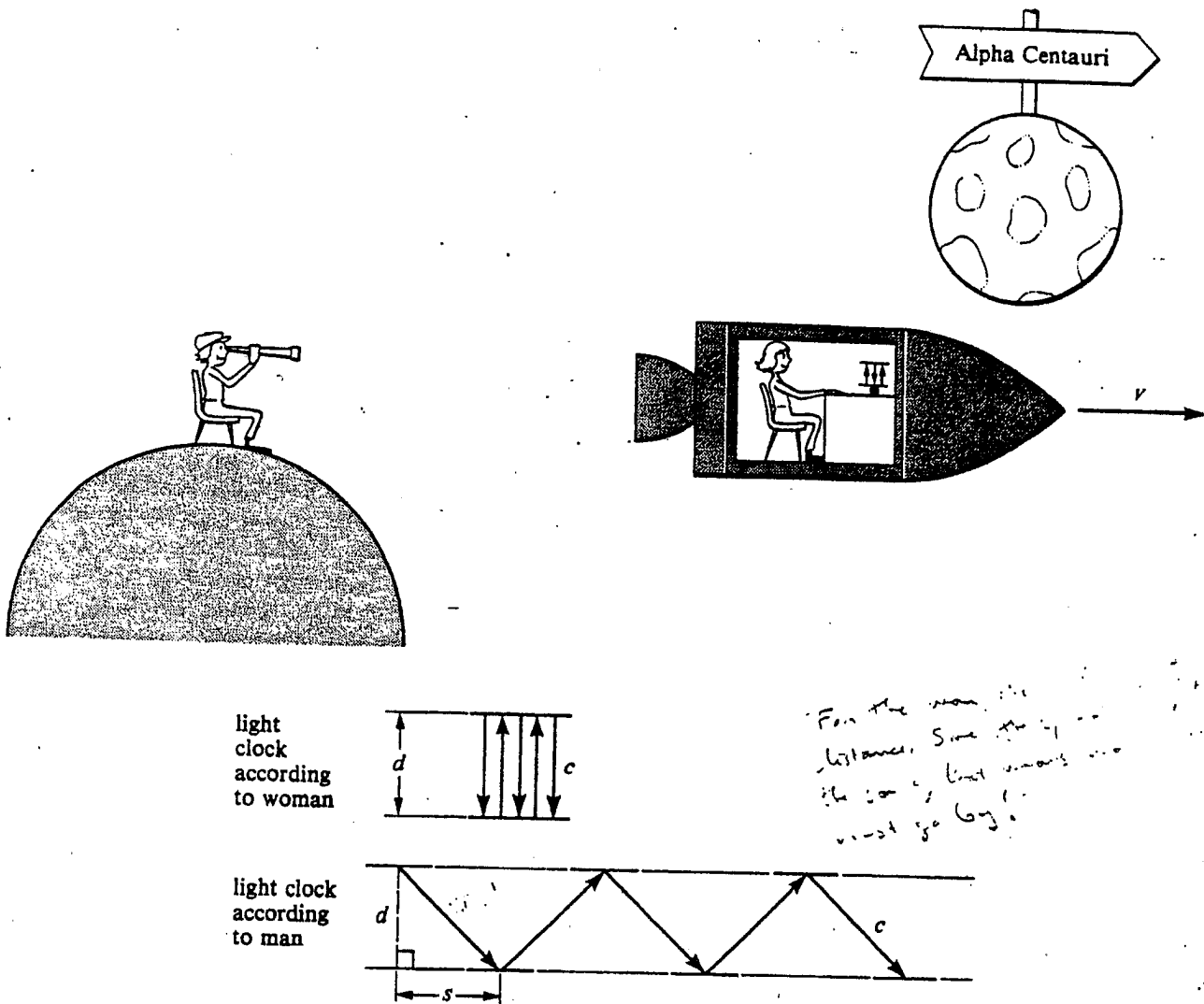


Figure 3.25. A rocket trip to Alpha Centauri at relativistic speeds provides an illustration of time dilation and the so-called "twin paradox."

Now, where is the paradox in this? Well, according to the principle of relativity, shouldn't the woman feel justified in thinking that it is she who is at rest and that it is the man (and the rest of the universe) who is moving? Thus, shouldn't she think that she ages more quickly than her brother because it is he who is moving? And so she does! However, there is still no paradox, because while she is moving there is no way they can get together to compare wrinkles. To compare themselves, she has to decelerate to a stop, turn around, accelerate back, decelerate, and come to a stop at the Earth. While decelerating and accelerating (with respect to the universe), she is no longer in an inertial frame, and we would have

to use a more complicated argument to figure out what is going on. But the net result would be that after she arrives back on Earth, she would indeed be younger than her twin brother! If this is confusing, just accept the result (3.11) and note that the woman's position is not identical to the man's, because it is she who is moving with respect to the "fixed" stars, not he.

Lorentz Contraction

Now, the persistent reader may argue that there is still something wrong. Isn't Alpha Centauri more than four light years away? How could the woman think she

Curved Space

note to students:
this very difficult

42-1 Curved spaces with two dimensions

According to Newton everything attracts everything else with a force inversely proportional to the square of the distance from it, and objects respond to forces with accelerations proportional to the forces. They are Newton's laws of universal gravitation and of motion. As you know, they account for the motions of balls, planets, satellites, galaxies, and so forth.

Einstein had a different interpretation of the law of gravitation. According to him, space and time—which must be put together as space-time—are *curved* near heavy masses. And it is the attempt of things to go along “straight lines” in this curved space-time which makes them move the way they do. Now that is a complex idea—very complex. It is the idea we want to explain in this chapter.

Our subject has three parts. One involves the effects of gravitation. Another involves the ideas of space-time which we already studied. The third involves the idea of curved space-time. We will simplify our subject in the beginning by not worrying about gravity and by leaving out the time—discussing just curved space. We will talk later about the other parts, but we will concentrate now on the idea of curved space—what is meant by curved space, and, more specifically, what is meant by curved space in this application of Einstein. Now even that much turns out to be somewhat difficult in three dimensions. So we will first reduce the problem still further and talk about what is meant by the words “curved space” in two dimensions.

In order to understand this idea of curved space in two dimensions you really have to appreciate the limited point of view of the character who lives in such a space. Suppose we imagine a bug with no eyes who lives on a plane, as shown in Fig. 42-1. He can move only on the plane, and he has no way of knowing that there is any way to discover any “outside world.” (He hasn't got your imagination.) We are, of course, going to argue by analogy. *We* live in a three-dimensional world, and we don't have any imagination about going off our three-dimensional world in a new direction; so we have to think the thing out by analogy. It is as though we were bugs living on a plane, and there was a space in another direction. That's why we will first work with the bug, remembering that he must live on his surface and can't get out.

As another example of a bug living in two dimensions, let's imagine one who lives on a sphere. We imagine that he can walk around on the surface of the sphere, as in Fig. 42-2 but that he can't look “up,” or “down,” or “out.”

Now we want to consider still a *third* kind of creature. He is also a bug like the others, and also lives on a plane, as our first bug did, but this time the plane is peculiar. The temperature is different at different places. Also, the bug and any rulers he uses are all made of the same material which expands when it is heated. Whenever he puts a ruler somewhere to measure something the ruler expands immediately to the proper length for the temperature at that place. Wherever he puts any object—himself, a ruler, a triangle, or anything—the thing stretches itself because of the thermal expansion. Everything is longer in the hot places than it is in the cold places, and everything has the same coefficient of expansion. We will call the home of our third bug a “hot plate,” although we will particularly want to think of a special kind of hot plate that is cold in the center and gets hotter as we go out toward the edges (Fig. 42-3).

Now we are going to imagine that our bugs begin to study geometry. Although we imagine that they are blind so that they can't see any “outside” world,

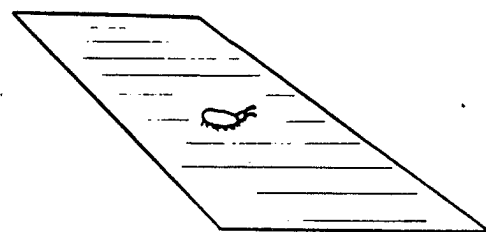
42-1 Curved spaces with two dimensions**42-2 Curvature in three-dimensional space****42-3 Our space is curved****42-4 Geometry in space-time****42-5 Gravity and the principle of equivalence****42-6 The speed of clocks in a gravitational field****42-7 The curvature of space-time****42-8 Motion in curved space-time****42-9 Einstein's theory of gravitation**

Fig. 42-1. A bug on a plane surface.

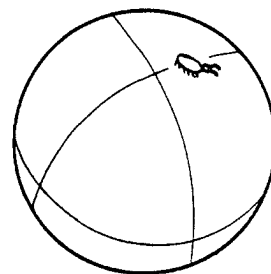


Fig. 42-2. A bug on a sphere.

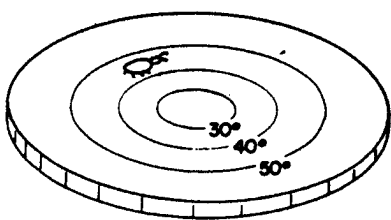


Fig. 42-3. A bug on a hot plate.

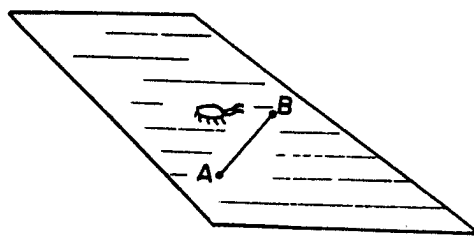


Fig. 42-4. Making a "straight line" on a plane.

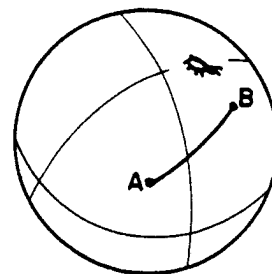


Fig. 42-5. Making a "straight line" on a sphere.

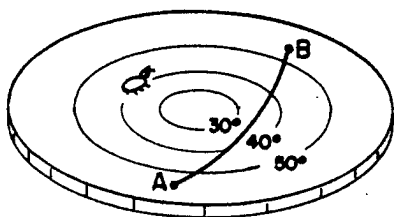


Fig. 42-6. Making a "straight line" on the hot plate.

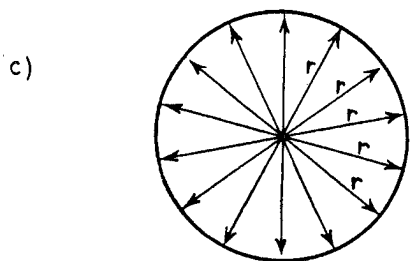
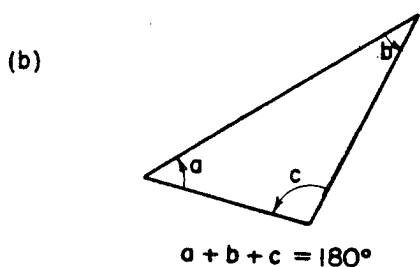
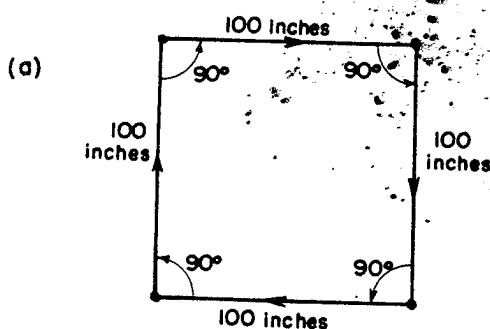


Fig. 42-7. A square, triangle, and circle in a flat space.

they can do a lot with their legs and feelers. They can draw lines, and they can make rulers, and measure off lengths. First, let's suppose that they start with the simplest idea in geometry. They learn how to make a straight line—defined as the shortest line between two points. Our first bug—see Fig. 42-4—learns to make very good lines. But what happens to the bug on the sphere? He draws his straight line as the shortest distance—for him—between two points, as in Fig. 42-5. It may look like a curve to us, but he has no way of getting off the sphere and finding out that there is "really" a shorter line. He just knows that if he tries any other path in his world it is always longer than his straight line. So we will let him have his straight line as the shortest arc between two points. (It is, of course an arc of a great circle.)

Finally, our third bug—the one in Fig. 42-3—will also draw "straight lines" that look like curves to us. For instance, the shortest distance between A and B in Fig. 42-6 would be on a curve like the one shown. Why? Because when his line curves out toward the warmer parts of his hot plate, the rulers get longer (from our omniscient point of view) and it takes fewer "yardsticks" laid end-to-end to get from A to B. So for him the line is straight—he has no way of knowing that there could be someone out in a strange three-dimensional world who would call a different line "straight."

We think you get the idea now that all the rest of the analysis will always be from the point of view of the creatures on the particular surfaces and not from our point of view. With that in mind let's see what the rest of their geometries looks like. Let's assume that the bugs have all learned how to make two lines intersect at right angles. (You can figure out how they could do it.) Then our first bug (the one on the normal plane) finds an interesting fact. If he starts at the point A and makes a line 100 inches long, then makes a right angle and marks off another 100 inches, then makes another right angle and goes another 100 inches, then makes a third right angle and a fourth line 100 inches long, he ends up right at the starting point as shown in Fig. 42-7(a). It is a property of his world—one of the facts of his "geometry."

Then he discovers another interesting thing. If he makes a triangle—a figure with three straight lines—the sum of the angles is equal to 180° , that is, to the sum of two right angles. See Fig. 42-7(b).

Then he invents the circle. What's a circle? A circle is made this way: You rush off on straight lines in many many directions from a single point, and lay out a lot of dots that are all the same distance from that point, and lay out a lot of dots that are all the same distance from that point. See Fig. 42-7(c). (We have to be careful how we define these things because we've got to be able to make the analogs for the other fellows.) Of course, its equivalent to the curve you can make by swinging a ruler around a point. Anyway, our bug learns how to make circles. Then one day he thinks of measuring the distance around a circle. He measures several circles and finds a neat relationship: The distance around is always the same number times the radius r (which is, of course, the distance from the center out to the curve). The circumference and the radius always have the same ratio—approximately 6.283—independent of the size of the circle.

Now let's see what our other bugs have been finding out about their geometries. First, what happens to the bug on the sphere when he tries to make a "square"?

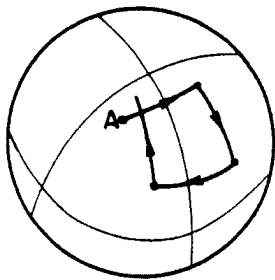


Fig. 42-8. Trying to make a "square" on a sphere.

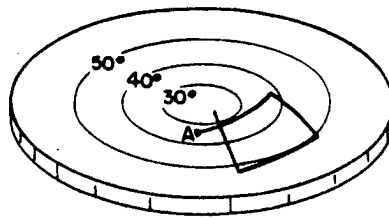


Fig. 42-9. Trying to make a "square" on the hot plate.

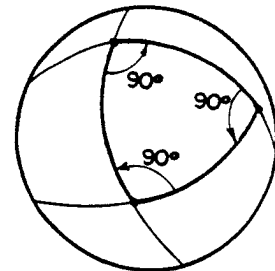


Fig. 42-10. On a sphere a "triangle" can have three 90° angles.

If he follows the prescription we gave above, he would probably think that the result was hardly worth the trouble. He gets a figure like the one shown in Fig. 42-8. His endpoint *B* isn't on top of the starting point *A*. It doesn't work out to a closed figure at all. Get a sphere and try it. A similar thing would happen to our friend on the hot plate. If he lays out four straight lines of equal length—as measured with his expanding rulers—joined by right angles he gets a picture like the one in Fig. 42-9.

Now suppose that our bugs had each had their own Euclid who had told them what geometry "should" be like, and that they had checked him out roughly by making crude measurements on a *small* scale. Then as they tried to make accurate squares on a larger scale they would discover that something was wrong. The point is, that just by *geometrical measurements* they would discover that something was the matter with their space. We define a *curved space* to be a space in which the geometry is not what we expect for a plane. The geometry of the bugs on the sphere or on the hot plate is the geometry of a curved space. The rules of Euclidian geometry fail. And it isn't necessary to be able to lift yourself out of the plane in order to find out that the world that you live in is curved. It isn't necessary to circumnavigate the globe in order to find out that it is a ball. You can find out that you live on a ball by laying out a square. If the square is very small you will need a lot of accuracy, but if the square is large the measurement can be done more crudely.

Let's take the case of a triangle on a plane. The sum of the angles is 180 degrees. Our friend on the sphere can find triangles that are very peculiar. He can, for example, find triangles which have *three right angles*. Yes indeed! One is shown in Fig. 42-10. Suppose our bug starts at the north pole and makes a straight line all the way down to the equator. Then he makes a right angle and another perfect straight line the same length. Then he does it again. For the very special length he has chosen he gets right back to his starting point, and also meets the first line with a right angle. So there is no doubt that for him this triangle has three right angles, or 270 degrees in the sum. It turns out that for him the sum of the angles of the triangle is *always* greater than 180 degrees. In fact, the excess (for the special case shown, the extra 90 degrees) is proportional to how much area the triangle has. If a triangle on a sphere is very small, its angles add up to very nearly 180 degrees, only a little bit over. As the triangle gets bigger the discrepancy goes up. The bugs on the hot plate would discover similar difficulties with their triangles.

Let's look next at what our other bugs find out about circles. They make circles and measure their circumferences. For example, the bug on the sphere might make a circle like the one shown in Fig. 42-11. And he would discover that the circumference is *less* than 2π times the radius. (You can see that because from the wisdom of our three-dimensional view it is obvious that what he calls the "radius" is a curve which is *longer* than the true radius of the circle.) Suppose that the bug on the sphere had read Euclid, and decided to predict a radius by dividing the circumference *C* by 2π , taking

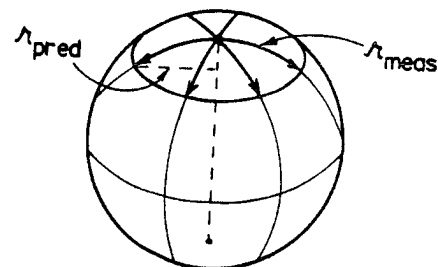


Fig. 42-11. Making a circle on a sphere.

$$r_{\text{pred}} = \frac{C}{2\pi} \quad (42.1)$$

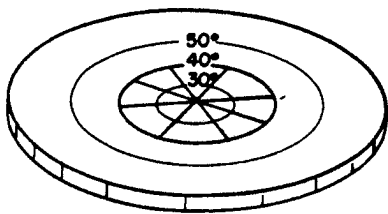


Fig. 42-12. Making a circle on the hot plate.

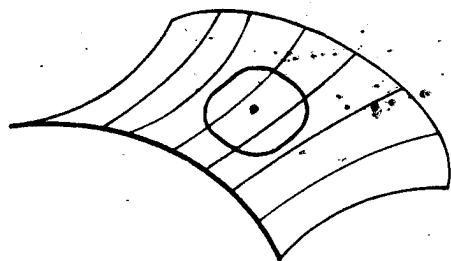


Fig. 42-13. A "circle" on a saddle-shaped surface.

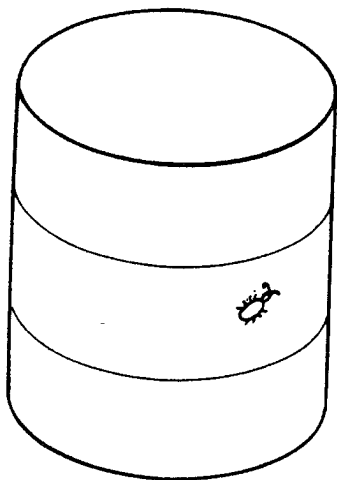


Fig. 42-14. A two-dimensional space with zero intrinsic curvature.

Then he would find that the measured radius was larger than the predicted radius. Pursuing the subject, he might define the difference to be the "excess radius," and write

$$r_{\text{meas}} - r_{\text{pred}} = r_{\text{excess}}, \quad (42.2)$$

and study how the excess radius effect depended on the size of the circle.

Our bug on the hot plate would discover a similar phenomenon. Suppose he was to draw a circle centered at the cold spot on the plate as in Fig. 42-12. If we were to watch him as he makes the circle we would notice that his rulers are short near the center and get longer as they are moved outward—although the bug doesn't know it, of course. When he measures the circumference the ruler is long all the time, so he, too, finds out that the measured radius is longer than the predicted radius, $C/2\pi$. The hot-plate bug also finds an "excess radius effect." And again the size of the effect depends on the radius of the circle.

We will define a "curved space" as one in which these types of geometrical errors occur: The sum of the angles of a triangle is different from 180 degrees; the circumference of a circle divided by 2π is not equal to the radius; the rule for making a square doesn't give a closed figure. You can think of others.

We have given two different examples of curved space: the sphere and the hot plate. But it is interesting that if we choose the right temperature variation as a function of distance on the hot plate, the two geometries will be exactly the same. It is rather amusing. We can make the bug on the hot plate get exactly the same answers as the bug on the ball. For those who like geometry and geometrical problems we'll tell you how it can be done. If you assume that the length of the rulers (as determined by the temperature) goes in proportion to one plus some constant times the square of the distance away from the origin, then you will find that the geometry of that hot plate is exactly the same in all details† as the geometry of the sphere.

There are, of course, other kinds of geometry. We could ask about the geometry of a bug who lived on a pear, namely something which has a sharper curvature in one place and a weaker curvature in the other place, so that the excess in angles in triangles is more severe when he makes little triangles in one part of his world than when he makes them in another part. In other words, the curvature of a space can vary from place to place. That's just a generalization of the idea. It can also be imitated by a suitable distribution of temperature on a hot plate.

We may also point out that the results could come out with the opposite kind of discrepancies. You could find out, for example, that all triangles when they are made too large have the sum of their angles *less* than 180 degrees. That may sound impossible, but it isn't at all. First of all, we could have a hot plate with the temperature decreasing with the distance from the center. Then all the effects would be reversed. But we can also do it purely geometrically by looking at the two-dimensional geometry of the surface of a saddle. Imagine a saddle-shaped surface like the one sketched in Fig. 42-13. Now draw a "circle" on the surface, defined as the locus of all points the same distance from a center. This circle is a curve that oscillates up and down with a scallop effect. So its circumference is larger than you would expect from calculating $2\pi r$. So $C/2\pi$ is now less than r . The "excess radius" would be negative.

Spheres and pears and such are all surfaces of *positive* curvatures; and the others are called surfaces of *negative* curvature. In general, a two-dimensional world will have a curvature which varies from place to place and may be positive in some places and negative in other places. In general, we mean by a curved space simply one in which the rules of Euclidean geometry break down with one sign of discrepancy or the other. The amount of curvature—defined, say, by the excess radius—may vary from place to place.

We might point out that, from our definition of curvature, a cylinder is, surprisingly enough, not curved. If a bug lived on a cylinder, as shown in Fig. 42-14, he would find out that triangles, squares, and circles would all have the

† Except for the one point at infinity.

same behavior they have on a plane. This is easy to see, by just thinking about how all the figures will look if the cylinder is unrolled onto a plane. Then all the geometrical figures can be made to correspond exactly to the way they are in a plane. So there is no way for a bug living on a cylinder (assuming that he doesn't go all the way around, but just makes local measurements) to discover that his space is curved. In our technical sense, then, we consider that his space is *not* curved. What we want to talk about is more precisely called *intrinsic* curvature; that is, a curvature which can be found by measurements only in a local region. (A cylinder has no intrinsic curvature.) This was the sense intended by Einstein when he said that our space is curved. But we as yet only have defined a curved space in two dimensions; we must go onward to see what the idea might mean in three dimensions.

42-2 Curvature in three-dimensional space

We live in three-dimensional space and we are going to consider the idea that three-dimensional space is curved. You say, "But how can you imagine it being bent in any direction?" Well, we can't imagine space being bent in any direction because our imagination isn't good enough. (Perhaps it's just as well that we can't imagine too much, so that we don't get too free of the real world.) But we can still *define* a curvature without getting out of our three-dimensional world. All we have been talking about in two dimensions was simply an exercise to show how we could get a definition of curvature which didn't require that we be able to "look in" from the outside.

We can determine whether our world is curved or not in a way quite analogous to the one used by the gentlemen who live on the sphere and on the hot plate. We may not be able to distinguish between two such cases but we certainly can distinguish those cases from the flat space, the ordinary plane. How? Easy enough: We lay out a triangle and measure the angles. Or we make a great big circle and measure the circumference and the radius. Or we try to lay out some accurate squares, or try to make a cube. In each case we test whether the laws of geometry work. If they don't work, we say that our space is curved. If we lay out a big triangle and the sum of its angles exceeds 180 degrees, we can say our space is curved. Or if the measured radius of a circle is not equal to its circumference over 2π , we can say our space is curved.

You will notice that in three dimensions the situation can be much more complicated than in two. At any one place in two dimensions there is a certain amount of curvature. But in three dimensions there can be *several components* to the curvature. If we lay out a triangle in some plane, we may get a different answer than if we orient the plane of the triangle in a different way. Or take the example of a circle. Suppose we draw a circle and measure the radius and it doesn't check with $C/2\pi$ so that there is some excess radius. Now we draw another circle at right angles—as in Fig. 42-15. There's no need for the excess to be exactly the same for both circles. In fact, there might be a positive excess for a circle in one plane, and a defect (negative excess) for a circle in the other plane.

Perhaps you are thinking of a better idea: Can't we get around all of these components by using a *sphere* in three dimensions? We can specify a sphere by taking all the points that are the same distance from a given point in space. Then we can measure the surface area by laying out a fine scale rectangular grid on the surface of the sphere and adding up all the bits of area. According to Euclid the total area A is supposed to be 4π times the square of the radius; so we can define a "predicted radius" as $\sqrt{A/4\pi}$. But we can also measure the radius directly by digging a hole to the center and measuring the distance. Again, we can take the measured radius minus the predicted radius and call the difference the radius excess,

$$r_{\text{excess}} = r_{\text{meas}} - \left(\frac{\text{measured area}}{4\pi} \right)^{1/2},$$

which would be a perfectly satisfactory measure of the curvature. It has the great advantage that it doesn't depend upon how we orient a triangle or a circle.

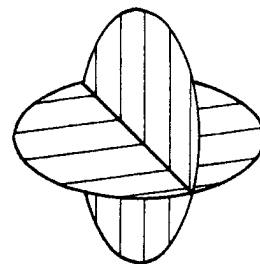


Fig. 42-15. The excess radius may be different for circles with different orientations.

But the excess radius of a sphere also has a disadvantage; it doesn't completely characterize the space. It gives what is called the *mean curvature* of the three-dimensional world, since there is an averaging effect over the various curvatures. Since it is an average, however, it does not solve completely the problem of defining the geometry. If you know only this number you can't predict all properties of the geometry of the space, because you can't tell what would happen with circles of different orientation. The complete definition requires the specification of six "curvature numbers" at each point. Of course the mathematicians know how to write all those numbers. You can read someday in a mathematics book how to write them all in a high-class and elegant form, but it is first a good idea to know in a rough way what it is that you are trying to write about. For most of our purposes the average curvature will be enough.†

42-3 Our space is curved

Now comes the main question. Is it true? That is, is the actual physical three-dimensional space we live in curved? Once we have enough imagination to realize the possibility that space might be curved, the human mind naturally gets curious about whether the real world is curved or not. People have made direct geometrical measurements to try to find out, and haven't found any deviations. On the other hand, by arguments about gravitation, Einstein discovered that space *is* curved, and we'd like to tell you what Einstein's law is for the amount of curvature, and also tell you a little bit about how he found out about it.

Einstein said that space is curved and that matter is the source of the curvature. (Matter is also the source of gravitation, so gravity is related to the curvature—but that will come later in the chapter.) Let us suppose, to make things a little easier, that the matter is distributed continuously with some density, which may vary, however, as much as you want from place to place.‡ The rule that Einstein gave for the curvature is the following: If there is a region of space with matter in it and we take a sphere small enough that the density ρ of matter inside it is effectively constant, then the radius excess for the sphere is proportional to the mass inside the sphere. Using the definition of excess radius, we have

$$\text{Radius excess} = \sqrt{\frac{A}{4\pi}} - r_{\text{meas}} = \frac{G}{3c^2} \cdot M. \quad (42.3)$$

Here, G is the gravitational constant (of Newton's theory), c is the velocity of light, and $M = 4\pi\rho r^3/3$ is the mass of the matter inside the sphere. This is Einstein's law for the mean curvature of space.

Suppose we take the earth as an example and forget that the density varies from point to point—so we won't have to do any integrals. Suppose we were to measure the surface of the earth very carefully, and then dig a hole to the center and measure the radius. From the surface area we could calculate the predicted radius we would get from setting the area equal to $4\pi r^2$. When we compared the predicted radius with the actual radius, we would find that the actual radius exceeded the predicted radius by the amount given in Eq. (42.3). The constant $G/3c^2$ is about 2.5×10^{-29} cm per gram, so for each gram of material the measured radius is off by 2.5×10^{-29} cm. Putting in the mass of the earth, which is about 6×10^{27} grams, it turns out that the earth has 1.5-millimeters more radius

† We should mention one additional point for completeness. If you want to carry the hot-plate model of curved space over into three dimensions you must imagine that the length of the ruler depends not only on where you put it, but also on which orientation the ruler has when it is laid down. It is a generalization of the simple case in which the length of the ruler depends on where it is, but is the same if set north-south, or east-west, or up-down. This generalization is needed if you want to represent a three-dimensional space with any arbitrary geometry with such a model, although it happens not to be necessary for two dimensions.

‡ Nobody—not even Einstein—knows how to do it if mass comes concentrated at points.

than it should have for its surface area.† Doing the same calculation for the sun, you find that the sun's radius is one-half a kilometer too long.

You should note that the law says that the average curvature above the surface area of the earth is zero. But that does not mean that all the components of the curvature are zero. There may still be—and, in fact, there is—some curvature above the earth. For a circle in a plane there will be an excess radius of one sign for some orientations and of the opposite sign for other orientations. It just turns out that the average over a sphere is zero when there is no mass inside it. Incidentally, it turns out that there is a relation between the various components of the curvature and the variation of the average curvature from place to place. So if you know the average curvature everywhere, you can figure out the details of the curvature at each place. The average curvature above the earth varies with altitude, so the space there is curved. And it is that curvature that we see as a gravitational force.

Suppose we have a bug on a plane, and suppose that the “plane” has little pimples in the surface. Wherever there is a pimple the bug would conclude that his space had little local regions of curvature. We have the same thing in three dimensions. Wherever there is a lump of matter, our three-dimensional space has a local curvature—a kind of three-dimensional pimple.

If we make a lot of bumps on a plane there might be an overall curvature besides all the pimples—the surface might become like a ball. It would be interesting to know whether our space has a net average curvature as well as the local pimples due to the lumps of matter like the earth and the sun. The astrophysicists have been trying to answer that question by making measurements of galaxies at very large distances. For example, if the number of galaxies we see in a spherical shell at a large distance is different from what we would expect from our knowledge of the radius of the shell, we would have a measure of the excess radius of a tremendously large sphere. From such measurements it is hoped to find out whether our whole universe is flat on the average, or round—whether it is “closed,” like a sphere, or “open” like a plane. You may have heard about the debates that are going on about this subject. There are debates because the astronomical measurements are still completely inconclusive; the experimental data are not precise enough to give a definite answer. Unfortunately, we don't have the slightest idea about the overall curvature of our universe on a large scale.

42-4 Geometry in space-time

Now we have to talk about time. As you know from the special theory of relativity, measurements of space and measurements of time are interrelated. And it would be kind of crazy to have something happening to the space, without the time being involved in the same thing. You will remember that the measurement of time depends on the speed at which you move. For instance, if we watch a guy going by in a spaceship we see that things happen more slowly for him than for us. Let's say he takes off on a trip and returns in 100 seconds flat by our watches; his watch might say that he had been gone for only 95 seconds. In comparison with ours, his watch—and all other processes, like his heart beat—have been running slow.

Now let's consider an interesting problem. Suppose you are the one in the spaceship. We ask you to start off at a given signal and return to your starting place just in time to catch a later signal—at, say, exactly 100 seconds later according to our clock. And you are also asked to make the trip in such a way that your watch will show the longest possible elapsed time. How should you move? You should stand still. If you move at all your watch will read less than 100 sec when you get back.

Suppose, however, we change the problem a little. Suppose we ask you to start at Point *A* on a given signal and go to point *B* (both fixed relative to us), and to do it in such a way that you arrive back just at the time of a second signal

† Approximately, because the density is not independent of radius as we are assuming.

(say 100 seconds later according to our fixed clock). Again you are asked to make the trip in the way that lets you arrive with the latest possible reading on your watch. How would you do it? For which path and schedule will *your* watch show the greatest elapsed time when you arrive? The answer is that you will spend the longest time from *your* point of view if you make the trip by going at a uniform speed along a straight line. Reason: Any extra motions and any extra-high speeds will make your clock go slower. (Since the time deviations depend on the square of the velocity, what you lose by going extra fast at one place you can never make up by going extra slowly in another place.)

The point of all this is that we can use the idea to define "a straight line" in space-time. The analog of a straight line in space is for space-time a *motion* at uniform velocity in a constant direction.

The curve of shortest distance in space corresponds in space-time not to the path of shortest time, but to the one of *longest* time, because of the funny things that happen to signs of the t -terms in relativity. "Straight-line" motion—the analog of "uniform velocity along a straight line"—is then that motion which takes a watch from one place at one time to another place at another time in the way that gives the longest time reading for the watch. This will be our definition for the analog of a straight line in space-time.

42-5 Gravity and the principle of equivalence

Now we are ready to discuss the laws of gravitation. Einstein was trying to generate a theory of gravitation that would fit with the relativity theory that he had developed earlier. He was struggling along until he latched onto one important principle which guided him into getting the correct laws. That principle is based on the idea that when a thing is falling freely everything inside it seems weightless. For example, a satellite in orbit is falling freely in the earth's gravity, and an astronaut in it feels weightless. This idea, when stated with greater precision, is called *Einstein's principle of equivalence*. It depends on the fact that all objects fall with exactly the same acceleration no matter what their mass, or what they are made of. If we have a spaceship that is "coasting"—so it's in a free fall—and there is a man inside, then the laws governing the fall of the man and the ship are the same. So if he puts himself in the middle of the ship he will stay there. He doesn't fall *with respect to the ship*. That's what we mean when we say he is "weightless."

Now suppose you are in a rocket ship which is accelerating. Accelerating with respect to what? Let's just say that its engines are on and generating a thrust so that it is not coasting in a free fall. Also imagine that you are way out in empty space so that there are practically no gravitational forces on the ship. If the ship is accelerating with "1g" you will be able to stand on the "floor" and will feel your normal weight. Also if you let go of a ball, it will "fall" toward the floor. Why? Because the ship is accelerating "upward," but the ball has no forces on it, so it will not accelerate; it will get left behind. Inside the ship the ball will appear to have a downward acceleration of "1g."

Now let's compare that with the situation in a spaceship sitting at rest on the surface of the earth. Everything is the same! You would be pressed toward the floor, a ball would fall with an acceleration of 1g, and so on. In fact, how could you tell inside a space ship whether you are sitting on the earth or are accelerating in free space? According to Einstein's equivalence principle there is no way to tell if you only make measurements of what happens to things inside!

To be strictly correct, that is true only for one point inside the ship. The gravitational field of the earth is not precisely uniform, so a freely falling ball has a slightly different acceleration at different places—the direction changes and the magnitude changes. But if we imagine a strictly uniform gravitational field, it is completely imitated in every respect by a system with a constant acceleration. That is the basis of the principle of equivalence.

42-6 The speed of clocks in a gravitational field

Now we want to use the principle of equivalence for figuring out a strange thing that happens in a gravitational field. We'll show you something that happens in a rocket ship which you probably wouldn't have expected to happen in a gravitational field. Suppose we put a clock at the "head" of the rocket ship—that is, at the "front" end—and we put another identical clock at the "tail," as in Fig. 42-16. Let's call the two clocks *A* and *B*. If we compare these two clocks when the ship is accelerating, the clock at the head seems to run fast relative to the one at the tail. To see that, imagine that the front clock emits a flash of light each second, and that you are sitting at the tail comparing the arrival of the light flashes with the ticks of clock *B*. Let's say that the rocket is in the position *a* of Fig. 42-17 when clock *A* emits a flash, and at the position *b* when the flash arrives at clock *B*. Later on the ship will be at position *c* when the clock *A* emits its next flash, and at position *d* when you see it arrive at clock *B*.

The first flash travels the distance L_1 and the second flash travels the shorter distance L_2 . It is a shorter distance because the ship is accelerating and has a higher speed at the time of the second flash. You can see, then, that if the two flashes were emitted from clock *A* one second apart, they would arrive at clock *B* with a separation somewhat less than one second, since the second flash doesn't spend as much time on the way. The same thing will also happen for all the later flashes. So if you were sitting in the tail you would conclude that clock *A* was running faster than clock *B*. If you were to do the same thing in reverse—letting clock *B* emit light and observing it at clock *A*—you would conclude that *B* was running slower than *A*. Everything fits together and there is nothing mysterious about it all.

But now let's think of the rocket ship at rest in the earth's gravity. *The same thing happens.* If you sit on the floor with one clock and watch another one which is sitting on a high shelf, it will appear to run faster than the one on the floor! You say, "But that is wrong. The times should be the same. With no acceleration there's no reason for the clocks to appear to be out of step." But they must if the principle of equivalence is right. And Einstein insisted that the principle was right, and went courageously and correctly ahead. He proposed that clocks at different places in a gravitational field must appear to run at different speeds. But if one always appears to be running at a different speed with respect to the other, then so far as the first is concerned the other is running at a different rate.

But now you see we have the analog for clocks of the hot ruler we were talking about earlier, when we had the bug on a hot plate. We imagined that rulers and bugs and everything changed lengths in the same way at various temperatures so they could never tell that their measuring sticks were changing as they moved around on the hot plate. It's the same with clocks in a gravitational field. Every clock we put at a higher level is seen to go faster. Heartbeats go faster, all processes run faster.

If they didn't you would be able to tell the difference between a gravitational field and an accelerating reference system. The idea that time can vary from place to place is a difficult one, but it is the idea Einstein used, and it is correct—believe it or not.

Using the principle of equivalence we can figure out how much the speed of a clock changes with height in a gravitational field. We just work out the apparent discrepancy between the two clocks in the accelerating rocket ship. The easiest way to do this is to use the result we found in Chapter 34 of Vol. 1 for the Doppler effect. There, we found—see Eq. (34.14)—that if v is the relative velocity of a source and a receiver, the received frequency ω is related to the emitted frequency ω_0 by

$$\omega = \omega_0 \frac{1 + v/c}{\sqrt{1 - v^2/c^2}} \quad (42.4)$$

Now if we think of the accelerating rocket ship in Fig. 42-17 the emitter and re-

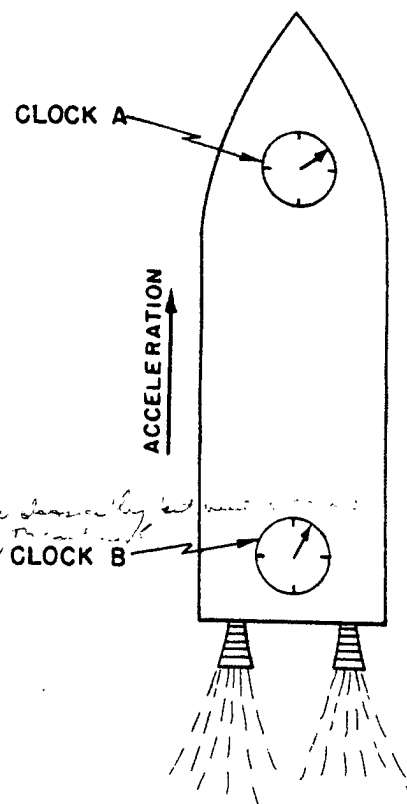


Fig. 42-16. An accelerating rocket ship with two clocks.

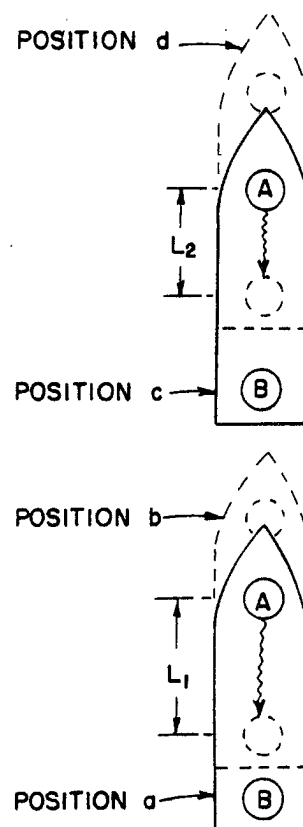


Fig. 42-17. A clock at the head of an accelerating rocket ship appears to run faster than a clock at the tail.

ceiver are moving with equal velocities at any one instant. But in the time that it takes the light signals to go from clock A to clock B the ship has accelerated. It has, in fact, picked up the additional velocity gt , where g is the acceleration and t is time it takes light to travel the distance H from A to B . This time is very nearly H/c . So when the signals arrive at B , the ship has increased its velocity by gH/c . The receiver always has this velocity with respect to the emitter at the instant the signal left it. So this is the velocity we should use in the Doppler shift formula, Eq. (42.4). Assuming that the acceleration and the length of the ship are small enough that this velocity is much smaller than c , we can neglect the term in v^2/c^2 . We have that

$$\omega = \omega_0 \left(1 + \frac{gH}{c^2} \right). \quad (42.5)$$

So for the two clocks in the spaceship we have the relation

$$(\text{Rate at the receiver}) = (\text{Rate of emission}) \left(1 + \frac{gH}{c^2} \right), \quad (42.6)$$

where H is the height of the emitter above the receiver.

From the equivalence principle the same result must hold for two clocks separated by the height H in a gravitational field with the free fall acceleration g .

This is such an important idea we would like to demonstrate that it also follows from another law of physics—from the conservation of energy. We know that the gravitational force on an object is proportional to its mass M , which is related to its total internal energy E by $M = E/c^2$. For instance, the masses of nuclei determined from the energies of nuclear reactions which transmute one nucleus into another agree with the masses obtained from atomic weights.

Now think of an atom which has a lowest energy state of total energy E_0 and a higher energy state E_1 , and which can go from the state E_1 to the state E_0 by emitting light. The frequency ω of the light will be given by

$$\hbar\omega = E_1 - E_0. \quad (42.7)$$

Now suppose we have such an atom in the state E_1 sitting on the floor, and we carry it from the floor to the height H . To do that we must do some work in carrying the mass $m_1 = E_1/c^2$ up against the gravitational force. The amount of work done is

$$\frac{E_1}{c^2} gH. \quad (42.8)$$

Then we let the atom emit a photon and go into the lower energy state E_0 . Afterward we carry the atom back to the floor. On the return trip the mass is E_0/c^2 ; we get back the energy

$$\frac{E_0}{c^2} gH, \quad (42.9)$$

so we have done a net amount of work equal to

$$\Delta U = \frac{E_1 - E_0}{c^2} gH. \quad (42.10)$$

When the atom emitted the photon it gave up the energy $E_1 - E_0$. Now suppose that the photon happened to go down to the floor and be absorbed. How much energy would it deliver there? You might at first think that it would deliver just the energy $E_1 - E_0$. But that can't be right if energy is conserved, as you can see from the following argument. We started with the energy E_1 at the floor. When we finish, the energy at the floor level is the energy E_0 of the atom in its lower state plus the energy E_{ph} received from the photon. In the meantime we have had to supply the additional energy ΔU of Eq. (42.10). If energy is conserved, the energy we end up with at the floor must be greater than we started with

by just the work we have done. Namely, we must have that

$$E_{ph} + E_0 = E_1 + \Delta U, \quad (42.11)$$

or

$$E_{ph} = (E_1 - E_0) + \Delta U.$$

It must be that the photon does *not* arrive at the floor with just the energy $E_1 - E_0$ it started with, but with a *little more energy*. Otherwise some energy would have been lost. If we substitute in Eq. (42.11) the ΔU we got in Eq. (42.10) we get that the photon arrives at the floor with the energy

$$E_{ph} = (E_1 - E_0) \left(1 + \frac{gH}{c^2} \right). \quad (42.12)$$

But a photon of energy E_{ph} has the frequency $\omega = E_{ph}/\hbar$. Calling the frequency of the *emitted* photon ω_0 —which is by Eq. (42.7) equal to $(E_1 - E_0)/\hbar$ —our result in Eq. (42.12) gives again the relation of (42.5) between the frequency of the photon when it is absorbed on the floor and the frequency with which it was emitted.

The same result can be obtained in still another way. A photon of frequency ω_0 has the energy $E_0 = \hbar\omega_0$. Since the energy E_0 has the gravitational mass E_0/c^2 the photon has a mass (*not* rest mass) $\hbar\omega_0/c^2$, and is “attracted” by the earth. In falling the distance H it will gain an additional energy $(\hbar\omega_0/c^2)gH$, so it arrives with the energy

$$E = \hbar\omega_0 \left(1 + \frac{gH}{c^2} \right).$$

But its frequency after the fall is E/\hbar , giving again the result in Eq. (42.5). Our ideas about relativity, quantum physics, and energy conservation all fit together only if Einstein's predictions about clocks in a gravitational field are right. The frequency changes we are talking about are normally very small. For instance, for an altitude difference of 20 meters at the earth's surface the frequency difference is only about two parts in 10^{15} . However, just such a change has recently been found experimentally using the Mössbauer effect.† Einstein was perfectly correct.

42-7 The curvature of space-time

Now we want to relate what we have just been talking about to the idea of curved space-time. We have already pointed out that if the time goes at different rates in different places, it is analogous to the curved space of the hot plate. But it is more than an analogy; it means that space-time *is* curved. Let's try to do some geometry in space-time. That may at first sound peculiar, but we have often made diagrams of space-time with distance plotted along one axis and time along the other. Suppose we try to make a rectangle in space-time. We begin by plotting a graph of height H versus t as in Fig. 42-18(a). To make the base of our rectangle we take an object which is *at rest* at the height H_1 and follow its world line for 100 seconds. We get the line BD in part (b) of the figure which is parallel to the t -axis. Now let's take another object which is 100 feet above the first one at $t = 0$. It starts at the point A in Fig. 42-18(c). Now we follow its world line for 100 seconds as measured by a clock at A . The object goes from A to C , as shown in part (d) of the figure. But notice that since time goes at a different rate at the two heights—we are assuming that there is a gravitational field—the two points C and D are not simultaneous. If we try to complete the square by drawing a line to the point C' which is 100 feet above D at the same time, as in Fig. 42-18(e), the pieces don't fit. And that's what we mean when we say that space-time is curved.

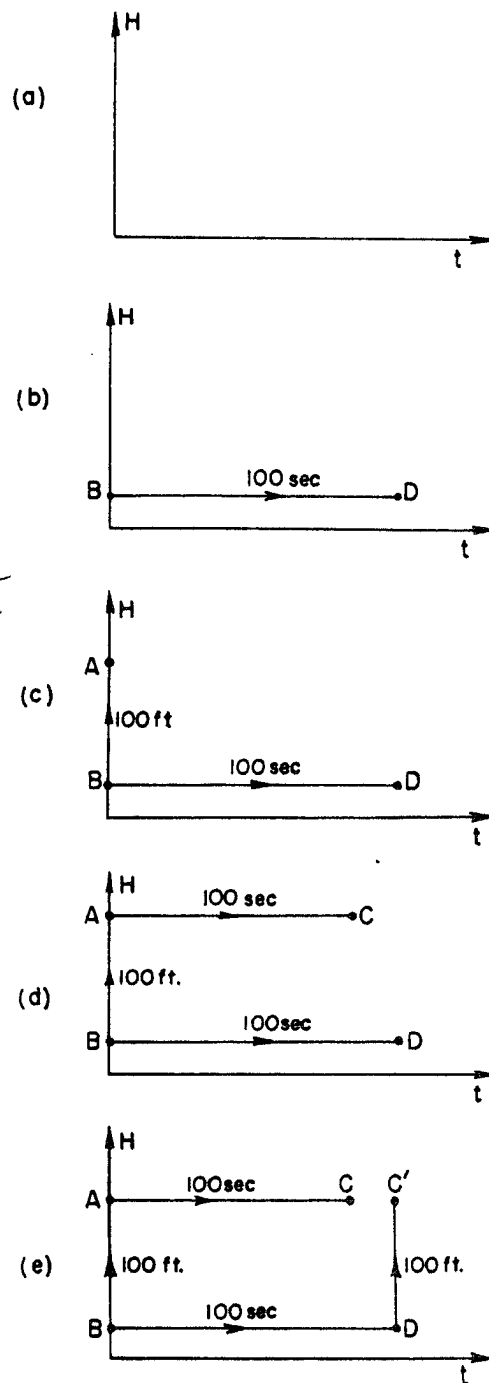


Fig. 42-18. Trying to make a rectangle in space-time.

† R. V. Pound and G. A. Rebka, Jr., *Physical Review Letters* Vol. 4, p. 337 (1960).

42-8 Motion in curved space-time

Let's consider an interesting little puzzle. We have two identical clocks, *A* and *B*, sitting together on the surface of the earth as in Fig. 42-19. Now we lift clock *A* to some height *H*, hold it there awhile, and return it to the ground so that it arrives at just the instant when clock *B* has advanced by 100 seconds. Then clock *A* will read something like 107 seconds, because it was running faster when it was up in the air. Now here is the puzzle. How should we move clock *A* so that it reads the latest possible time—always assuming that it returns when *B* reads 100 seconds? You say, "That's easy. Just take *A* as high as you can. Then it will run as fast as possible, and be the latest when you return." Wrong. You forgot something—we've only got 100 seconds to go up and back. If we go very high, we have to go very fast to get there and back in 100 seconds. And you mustn't forget the effect of special relativity which causes moving clocks to *slow down* by the factor $\sqrt{1 - v^2/c^2}$. This relativity effect works in the direction of making clock *A* read *less time* than clock *B*. You see that we have a kind of game. If we stand still with clock *A* we get 100 seconds. If we go up slowly to a small height and come down slowly we can get a little more than 100 seconds. If we go a little higher, maybe we can gain a little more. But if we go too high we have to move fast to get there, and we may slow down the clock enough that we end up with less than 100 seconds. What program of height versus time—how high to go and with what speed to get there, carefully adjusted to bring us back to clock *B* when it has increased by 100 seconds—will give us the largest possible time reading on clock *A*?

Answer: Find out how fast you have to throw a ball up into the air so that it will fall back to earth in exactly 100 seconds. The ball's motion—rising fast, slowing down, stopping, and coming back down—is exactly the right motion to make the time the maximum on a wrist watch strapped to the ball.

Now consider a slightly different game. We have two points *A* and *B* both on the earth's surface at some distance from one another. We play the same game that we did earlier to find what we call the straight line. We ask how we should go from *A* to *B* so that the time on our moving watch will be the longest—assuming we start at *A* on a given signal and arrive at *B* on another signal at *B* which we will say is 100 seconds later by a fixed clock. Now you say, "Well we found out before that the thing to do is to coast along a straight line at a uniform speed chosen so that we arrive at *B* exactly 100 seconds later. If we don't go along a straight line it takes more speed, and our watch is slowed down." But wait! That was before we took gravity into account. Isn't it better to curve upward a little bit and then come down? Then during part of the time we are higher up and our watch will run a little faster? It is, indeed. If you solve the mathematical problem of adjusting the curve of the motion so that the elapsed time of the moving watch is the most it can possibly be, you will find that the motion is a parabola—the same curve followed by something that moves on a free ballistic path in the gravitational field, as in Fig. 42-19. Therefore the law of motion in a gravitational field can also be stated: *An object always moves from one place to another so that a clock carried on it gives a longer time than it would on any other possible trajectory*—with, of course, the same starting and finishing conditions. The time measured by a moving clock is often called its "proper time." In free fall, the trajectory makes the proper time of an object a maximum.

Let's see how this all works out. We begin with Eq. (42.5) which says that the excess rate of the moving watch is

$$\frac{\omega_0 g H}{c^2} \quad (42.13)$$

Besides this, we have to remember that there is a correction of the opposite sign for the speed. For this effect we know that

$$\omega = \omega_0 \sqrt{1 - v^2/c^2}.$$

Although the principle is valid for any speed, we take an example in which the

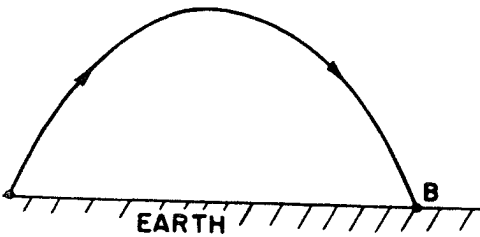


Fig. 42-19. In a uniform gravitational field the trajectory with the maximum proper time for a fixed elapsed time is a parabola.

speeds are always much less than c . Then we can write this equation as

$$\omega = \omega_0(1 - v^2/2c^2),$$

and the defect in the rate of our clock is

$$-\omega_0 \frac{v^2}{2c^2}. \quad (42.14)$$

Combining the two terms in (42.13) and (42.14) we have that

$$\Delta\omega = \frac{\omega_0}{c^2} \left(gH - \frac{v^2}{2} \right). \quad (42.15)$$

Such a frequency shift of our moving clock means that if we measure a time dt on a fixed clock, the moving clock will register the time

$$dt \left[1 + \left(\frac{gH}{c^2} - \frac{v^2}{2c^2} \right) \right], \quad (42.16)$$

The total time excess over the trajectory is the integral of the extra term with respect to time, namely

$$\frac{1}{c^2} \int \left(gH - \frac{v^2}{2} \right) dt, \quad (42.17)$$

which is supposed to be a maximum.

The term gH is just the gravitational potential ϕ . Suppose we multiply the whole thing by a constant factor $-mc^2$, where m is the mass of the object. The constant won't change the condition for the maximum, but the minus sign will just change the maximum to a minimum. Equation (42.16) then says that the object will move so that

$$\int \left(\frac{mv^2}{2} - m\phi \right) dt = \text{a minimum}. \quad (42.18)$$

But now the integrand is just the difference of the kinetic and potential energies. And if you look in Chapter 19 of Volume II you will see that when we discussed the principle of least action we showed that Newton's laws for an object in any potential could be written exactly in the form of Eq. (42.18).

42-9 Einstein's theory of gravitation

Einstein's form of the equations of motion—that the proper time should be a maximum in curved space-time—gives the same results as Newton's laws for low velocities. As he was circling around the earth, Gordon Cooper's watch was reading later than it would have in any other path you could have imagined for his satellite.†

So the law of gravitation can be stated in terms of the ideas of the geometry of space-time in this remarkable way. The particles always take the longest proper time—in space-time a quantity analogous to the "shortest distance." That's the law of motion in a gravitational field. The great advantage of putting it this way is that the law doesn't depend on any coordinates, or any other way of defining the situation.

Now let's summarize what we have done. We have given you two laws for gravity:

- (1) How the geometry of space-time changes when matter is present—namely, that the curvature expressed in terms of the excess radius is proportional to the mass inside a sphere, Eq. (42.3).

† Strictly speaking it is only a *local* maximum. We should have said that the proper time is larger than for any *nearby* path. For example, the proper time on an elliptical orbit around the earth need not be longer than on a ballistic path of an object which is shot to a great height and falls back down.

- (2) How objects move if there are only gravitational forces—namely, that objects move so that their proper time between two end conditions is a maximum.

Those two laws correspond to similar pairs of laws we have seen earlier. We originally described motion in a gravitational field in terms of Newton's inverse square law of gravitation and his laws of motion. Now laws (1) and (2) take their places. Our new pair of laws also correspond to what we have seen in electrodynamics. There we had our law—the set of Maxwell's equations—which determines the fields produced by charges. It tells how the character of "space" is changed by the presence of charged matter, which is what law (1) does for gravity. In addition, we had a law about how particles move in the given fields— $d(mv)/dt = q(E + v \times B)$. This, for gravity, is done by law (2).

In the laws (1) and (2) you have a precise statement of Einstein's theory of gravitation—although you will usually find it stated in a more complicated mathematical form. We should, however, make one further addition. Just as time scales change from place to place in a gravitational field, so do also the length scales. Rulers change lengths as you move around. It is impossible with space and time so intimately mixed to have something happen with time that isn't in some way reflected in space. Take even the simplest example: You are riding past the earth. What is "time" from *your* point of view is partly space from *our* point of view. So there must also be changes in space. It is the entire *space-time* which is distorted by the presence of matter, and this is more complicated than a change only in time scale. However, the rule that we gave in Eq. (42-3) is enough to determine completely all the laws of gravitation, provided that it is understood that this rule about the curvature of space applies not only from one man's point of view but is true for everybody. Somebody riding by a mass of material sees a different mass content because of the kinetic energy he calculates for its motion past him, and he must include the mass corresponding to that energy. The theory must be arranged so that everybody—no matter how he moves—will, when he draws a sphere, find that the excess radius is $G/3c^2$ times the total mass (or, better, $G/3c^4$ times the total energy content) inside the sphere. That this law—law (1)—should be true in any moving system is one of the great laws of gravitation, called *Einstein's field equation*. The other great law is (2)—that things must move so that the proper time is a maximum—and is called *Einstein's equation of motion*.

To write these laws in a complete algebraic form, to compare them with Newton's laws, or to relate them to electrodynamics is difficult mathematically. But it is the way our most complete laws of the physics of gravity look today.

Although they gave a result in agreement with Newton's mechanics for the simple example we considered, they do not always do so. The three discrepancies first derived by Einstein have been experimentally confirmed: The orbit of Mercury is not a fixed ellipse; starlight passing near the sun is deflected twice as much as you would think; and the rates of clocks depend on their location in a gravitational field. Whenever the predictions of Einstein have been found to differ from the ideas of Newtonian mechanics Nature has chosen Einstein's.

Let's summarize everything that we have said in the following way. First, time and distance rates depend on the place in space you measure them and on the time. This is equivalent to the statement that space-time is curved. From the measured area of a sphere we can define a predicted radius, $\sqrt{A/4\pi}$, but the actual measured radius will have an excess over this which is proportional (the constant is G/c^2) to the total mass contained inside the sphere. This fixes the exact degree of the curvature of space-time. And the curvature must be the same no matter who is looking at the matter or how it is moving. Second, particles move on "straight lines" (trajectories of maximum proper time) in this curved space-time. This is the content of Einstein's formulation of the laws of gravitation.

Curved space according to Frank Shur
note to students;
this very difficult

Black Holes

We saw that the escape speed from the surface of a neutron star was typically a healthy fraction of the speed of light. What radius R would an object of mass M have to have if the escape speed from its surface were equal to the speed of light? If we naively apply formula (7.3), set v_e equal to c , and solve the result for R , we obtain the desired radius—the *Schwarzschild radius*—as the expression

$$R_{\text{Sch}} = \frac{2GM}{c^2}. \quad (7.4)$$

Apart from quantum-mechanical effects, any object whose radius R becomes smaller than its Schwarzschild radius is doomed to collapse to a single point. No known force in nature can resist this collapse to singularity with an infinite density. The numerical value of the Schwarzschild radius for a $1M_{\odot}$ black hole (if there is a way to form them) is

$$R_{\text{Sch}} = 3 \times 10^5 \text{ cm for } M = 1M_{\odot},$$

and it is proportionally larger for larger values of M . Hence, stellar-mass black holes are only a few times smaller than neutron stars of comparable mass.

Shur, The Physical Universe

The expression $2GM/c^2$ which appears on the right-hand side of equation (7.4) was first derived by the nineteenth-century physicist Pierre Laplace. His derivation, which we have reproduced, contains two errors of assumption. The first error is to suppose that the kinetic energy of light is given by $mv^2/2$. It is not. The second error is to suppose that Newton's law of gravitation continues to hold when even light finds it difficult to escape the gravitational clutch of a massive body. It does not. Karl Schwarzschild was the first person to give the correct derivation for the size of an object—later named a **black hole** by John A. Wheeler—which would not let even light escape from its gravitational pull. For this reason, the expression $2GM/c^2$ is called the "Schwarzschild radius." The two errors of Laplace's derivation happen to cancel to give the correct-looking formula.

Schwarzschild's analysis shows that a photon which starts with wavelength λ_0 at position r outside a spherical gravitating mass M will have a longer wavelength λ when it emerges at infinity. The relation between λ and λ_0 is given by the **gravitational redshift** formula:

$$\frac{\lambda}{\lambda_0} = \left(1 - \frac{2GM}{c^2 r}\right)^{-1/2} \quad (7.5)$$

Notice that if $r = 2GM/c^2$, the above formula predicts that the photon will be redshifted to infinite wavelength λ and zero energy hc/λ . In other words, if a body is smaller than its Schwarzschild radius, all photons which tried to leave its surface would be redshifted into non-existence before they could reach infinity. This is the ultimate meaning of the Schwarzschild radius (7.4).

The correct derivation of equation (7.5) requires the use of Einstein's improvement of Newton's theory of gravitation. Einstein's theory of gravitation is, of course, general relativity. General relativity yields Newtonian gravity as an acceptable approximation when the gravitational field is relatively weak. For example, suppose in equation (7.5) we are at a value of r which is much greater than the Schwarzschild radius, $2GM/c^2$. Then, if we take the reciprocal of equation (7.5) and expand the resulting right-hand side by the binomial theorem for $2GM/c^2 r \ll 1$, we obtain the equation

$$\frac{\lambda_0}{\lambda} = 1 - \frac{GM}{c^2 r}$$

Multiplying by the energy $E = hc/\lambda_0$ of the photon at position r and identifying $m = E/c^2$ as the associated "gravitating mass" of the photon, we obtain

$$\frac{hc}{\lambda} = \frac{hc}{\lambda_0} - \frac{GMm}{r} \quad (7.6)$$

The physical interpretation of equation (7.6) in terms of the Newtonian principle of conservation of energy is

obvious. It states that the kinetic energy of the photon at infinity is equal to the sum of the kinetic and potential energies of the photon at r . The result is non-Newtonian, because we use quantum mechanics to identify hc/λ as the kinetic energy of a photon, and Einstein's equivalence between mass and energy $E = mc^2$ to identify $m = h/c\lambda$ as the "gravitating mass" of a photon.

We should not think, however, that Einstein's theory of gravity represents merely a quantitative modification of Newton's theory. General relativity gives quantitatively and qualitatively very different results from Newtonian theory when the escape speeds of matter from any gravitating mass approach the speed of light, as they must by definition near the Schwarzschild radius of a black hole. The most important qualitative distinction is the geometric interpretation of spacetime, r and t .

Gravitational Distortion of Spacetime

Spacetime is very distorted from that appropriate to special relativity in the vicinity of a black hole. Indeed, to be precise, we should not even think of $2GM/c^2$ as the "radius" of the black hole, since there is no operational way for us to measure such a "radius." If we tried to lay a tape measure across the black hole, the tape measure would be swallowed by the black hole. If we try to send light across the black hole and bounce it back with a mirror to time the round trip, the photons would fall into the black hole and never emerge from the other side. All such attempts to measure the "radius" of a black hole are doomed to fail.

Something we can measure is the circumference of the **event horizon** (Figure 7.8). The event horizon is the surface where outwardly traveling photons would barely be able to escape to infinity. If we are suitably suspended, we could imagine walking around the spherically symmetric event horizon of a Schwarzschild black hole and measuring the circumference. In this way, we would obtain a distance $4\pi GM/c^2$, where M is the mass obtained by applying Newton's laws to a small body orbiting the black hole at a very large distance. Defining the circumference measured this way, $4\pi GM/c^2$, to be $2\pi R_{\text{Sch}}$, we would then recover equation (7.4). In a similar fashion, we could imagine lowering many sheets of graph paper toward the event horizon and counting up how many sheets are needed to cover the latter. The surface area measured in this way would turn out to be $16\pi G^2 M^2/c^4$, i.e., $4\pi R_{\text{Sch}}^2$ with our definition of R_{Sch} .

The recovery of the usual Euclidean relationship between the circumference of a circle and the surface area of a sphere is a consequence of the angular symmetry of the problem. The *radial* direction suffers distortions from our usual geometric perceptions. To experience this distortion, imagine conducting the following surveying experiment. First, place yourself in a

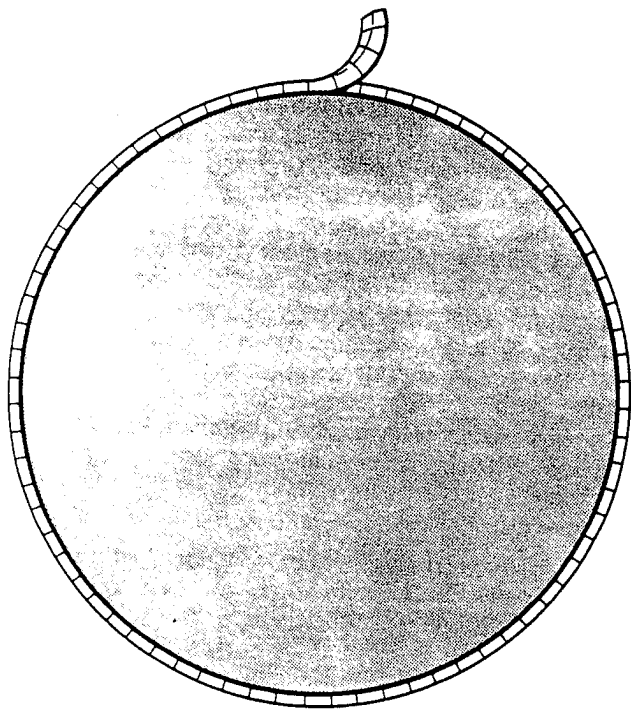


Figure 7.8. Although it is not possible to measure the “radius” of a black hole, one can imagine measuring its circumference by encircling its event horizon with a tape measure. The circumference obtained this way is equal to $2\pi R_{\text{Sch}}$, where $R_{\text{Sch}} = 2GM/c^2$, and M is the “gravitational mass” of the black hole.

powerful rocket ship at a great distance from the center of the black hole whose Schwarzschild “radius” is 3 km. Next, fly in a circle about it. What is the circumference? Two π times 30 km, you reply. At this point, you might feel justified to deduce that you are 30 km from the black hole, but don’t jump to conclusions. Try flying 10.67 km toward the black hole. Now, fly in a circle around it. What is the circumference? Two π times 20 km, you reply. Have we made a mistake? Shouldn’t you be 30 km – 10.67 km = 19.33 km, not 20 km, from the black hole? Well, try flying 21.92 km from your original starting point toward the black hole. Now, fly in a circle around it again. What is the circumference? Two π times 10 km, you reply. Now, you begin to become alarmed because Euclidean geometry predicts you should be 30 km – 21.92 km = 8.08 km, not 10 km, from the black hole. What’s going on? Try flying 28.52 km from your original starting point toward the black hole. You protest, won’t that take you inside the event horizon of the black hole? If the Schwarzschild “radius” is 3 km, isn’t it only 27 km from your original starting point? Don’t worry; try flying 28.32 km inwards. See, you’re OK; you can still fly in a circle around it. What is the

circumference? Two π times 5 km, you reply, badly shaken. Clearly, the strong gravitational field of the black hole must be distorting the usual spatial relationships. Figure 7.9 summarizes our thought-experiment findings in the form of an “embedding diagram.”

Because of the spherical symmetry, flight circles of arbitrary inclination around the black hole are equivalent. Let us plot only one angle of inclination. If the geometry of space had been Euclidean, we could have made a “polar plot” of the relationship between radial distance and circumference on a flat sheet of paper (Figure 7.9a). For our thought-experiment findings, we actually need a curved sheet of paper (Figure 7.9b). In Figure 7.9b the direction perpendicular to the curved paper has no physical meaning. The two directions on the surface of the curved paper correspond to physical directions in a single plane. Remember that the physical third dimension (the direction perpendicular to our physical plane) has been left off our “polar plot.” The “curvature” of our two-dimensional plot embedded in a third plotting dimension is purely a mathematical device constructed to allow us to visualize geometrically the distortion produced in space by the gravitational field of the black hole. If we had insisted on plotting a plane on a flat sheet of paper, one cm near the center would represent a different distance than one cm further out. The “curvature” of actual three-dimensional space cannot be visualized by us three-dimensional creatures; we need additional (artificial) spatial dimensions to plot three-dimensional space curvature. Differential geometry allows general relativists, however, to calculate symbolically the three-dimensional space curvature that they cannot visualize geometrically. The symmetry of the present problem allows us to throw away one of the actual spatial dimensions and to use the “embedding diagram” of Figure 7.9b. One glance at this embedding diagram suffices to show that you have to go a greater radial distance inward than Euclidean geometry (Figure 7.9a) would predict to measure a given smaller circumference. At great distances from the black hole, the distortions from Euclidean geometry are negligibly small, but the distortions become severe near the event horizon. In particular, the outsider’s notion of space ends with the event horizon, because the “curvature” associated with the gravity of the black hole is a curvature of spacetime, and not just a curvature of space alone. Space and time change roles across the event horizon.

To explore the curvature of spacetime near a black hole, let us imagine the following thought experiment. Let us imagine lowering you very carefully by a very strong cable toward the event horizon of a black hole. What do you see?

Well, when you are still far from the event horizon, you can see the constellations much as they usually appear in that part of the universe. As you approach closer to the event horizon, the positions of the stars in

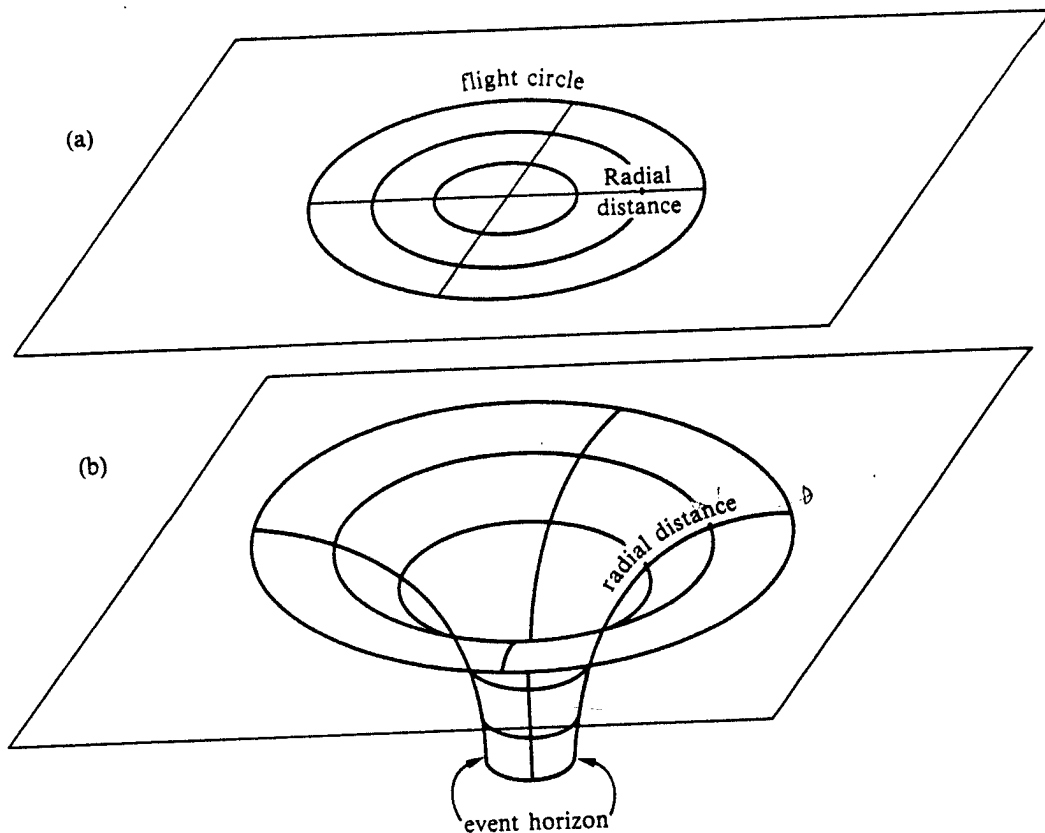


Figure 7.9. Flight circles in a plane (a) in Euclidean geometry, (b) in the non-Euclidean geometry near the event horizon of a black hole. In case (b), one has to travel a greater distance inward than in case (a) to have a flight circle of given smaller circumference. The radial direction in both cases is as indicated. At great distances from the event horizon (not drawn), the "curvature" of our embedding diagram becomes negligibly small, and the flight circles of case (b) have nearly the same geometry as case (a).

the constellations shift and distort. Do you begin to get a sinking feeling? When you are lowered to where the circumference is 1.5 times the circumference of the event horizon, you can look perpendicular to the black hole and see the back of your own head! Don't look toward the black hole. You can't see anything in that direction. It's completely dark toward the black hole. Look straight ahead, *perpendicular* to the black hole. You can now see the back of your own head because the photons leaving the back of your own head can orbit around the black hole to strike your eyes directly (Figure 7.10). It's a nice trick, and it's done with gravity, not with mirrors.

Now, I continue to lower you very carefully nearer and nearer to the event horizon. Careful! Watch out! With a sickening snap, the cable breaks, and you fall irretrievably toward the black hole. It's no use blaming me; I used the

strongest cable money can buy. The problem is that as you approach the event horizon, no force in nature—not the electromagnetic force, not the strong force, and certainly not the weak force—can prevent you from falling toward the black hole (Problem 7.8). As you fall, it appears to me—stationed comfortably at a safe distance from the black hole—that it takes you formally an infinite time to reach the event horizon. Nevertheless, once you start falling, effective communication ends between us in about a millisecond, because all photon signals you may frantically try to send out to me either are gravitationally redshifted to undetectably long wavelengths by the time they reach me, or fall with you into the black hole (Figure 7.11). In a sense, space and time seem to reverse roles as you approach the event horizon of the black hole. When you were far from the black hole,

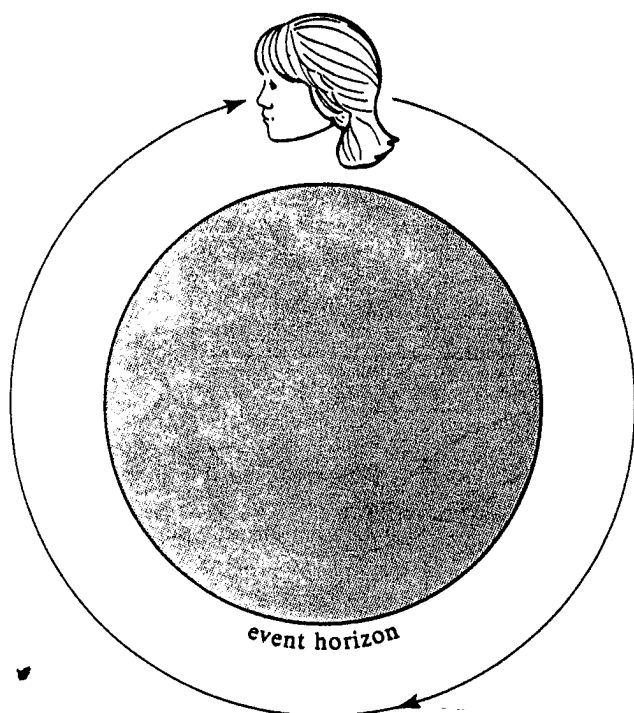


Figure 7.10. When at a circumference equal to 1.5 times the circumference of the event horizon of a black hole, a suitably suspended astronaut can see the back of her own head without the benefit of any mirrors.

by exerting enough effort, I could hold your position steady with respect to the black hole. But there was nothing that I could do to stop the forward march of time; I could not, for example, stop you from aging. As you fall toward the event horizon of the black hole, there is nothing I can do to stop you from falling, but time seems to have stopped moving for you!

Problem 7.8. The acceleration g of a freely falling body near a Schwarzschild black hole is given by the formula

$$g = \frac{GM/r^2}{(1 - 2GM/c^2r)^{1/2}},$$

when at a radius where $2\pi r$ is the circumference. Show that far from the event horizon, the above formula reduces to the usual Newtonian expression. Show also that when $r = R_{\text{sch}}$, the gravitational force becomes infinite, and therefore irresistible. An infinite quantity is not very convenient to work with; hence, it is convenient to define the "surface gravity" which is just the

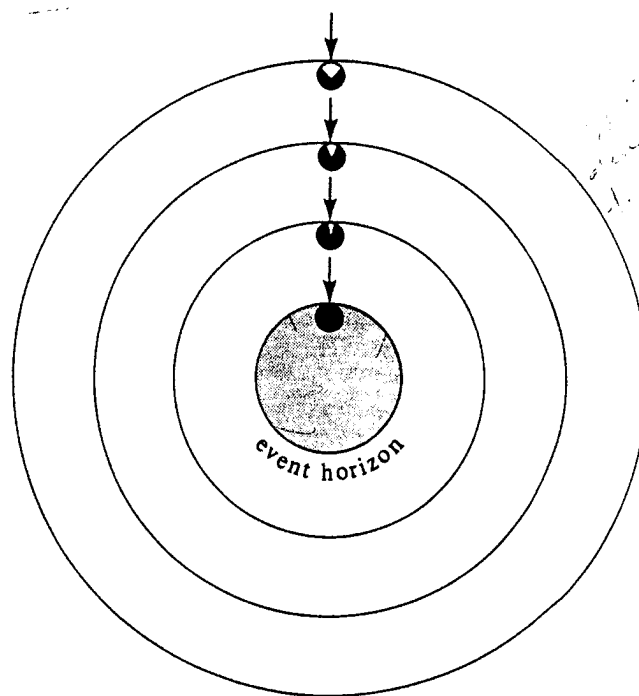


Figure 7.11. Journey to the center of a black hole. As an adventuresome explorer falls toward a black hole, he sends out light signals in all directions. The explorer sees locally the photons race away from him in all directions at the speed of light c (small circles). However, only the photons traveling in the outward direction of the light cone will escape from the gravitational clutch of the black hole to reach eventually any outside observer. The rest of the photons (dark part of small circle) are doomed to fall into the black hole with the explorer. The escape cone becomes smaller and smaller as the explorer approaches the event horizon. In particular, when he reaches it, only photons going exactly radially outward can escape.

numerator of the preceding formula evaluated at the event horizon, i.e.,

$$\text{"surface gravity"} = GM/R_{\text{sch}}^2 = c^4/4GM.$$

A similar quantity can be defined for rotating black holes, which do not possess complete angular symmetry, but the "surface gravity" also works out to a constant (depending only on the mass M and angular momentum J).

Let Δt_∞ be the time interval elapsed for an observer at infinity. Corresponding to this, the time elapsed for a fixed local observer is

$$\Delta t_0 = (1 - 2GM/c^2r)^{1/2} \Delta t_\infty.$$

Show that a local observer far from the black hole ages at the same rate as the observer at infinity, and that the former's aging rate stops in comparison with the latter's when the event horizon is approached.

A particle falling with a parabolic velocity (zero binding energy) takes time (observed by the falling particle)

$$\Delta t = \frac{2}{3} \left[\frac{r}{c} \left(\frac{r}{R_{\text{Sch}}} \right)^{1/2} - \frac{R_{\text{Sch}}}{c} \right]$$

to fall from a "radius" r to the event horizon. Calculate Δt numerically if $R_{\text{Sch}} = 3$ km and $r = 10$ km. The observer at infinity thinks that it takes an infinite time for the particle to reach the event horizon.

How do you react as you fall? According to your watch, it takes only a finite time to reach and cross the event horizon. Indeed, it takes you only a finite time to reach the very center of the black hole, where a singular catastrophe awaits you. Nothing very peculiar happens as you cross the event horizon, except that all photons which you send outward using your flashlight cannot escape the clutches of the black hole either. Well, actually, one peculiar thing does happen to you as you fall that I forgot to mention before sending you on this journey. As you fall feet first toward the center of the black hole, the gravitational pull on your feet is considerably stronger than the gravitational pull on your head. This "tidal force" soon stretches you out in a long long string. Sorry. As a consolation, if you survive this modern-day torture rack, as you cross the event horizon, you will have entered another "universe" whose properties (apart from mass and rotation) none of us on the "outside" will ever sample.

Thermodynamics of Black Holes

The above discussion of black holes ignores some interesting quantum-mechanical effects discovered only recently. On page 135 we noted that the surface area of a nonrotating black hole is proportional to the square of the mass M of the black hole. Classically, matter-energy can only fall into black holes, never fall out; so the mass of a nonrotating black hole can only increase, never decrease. This result holds also for rotating black holes, providing we replace the word "mass" by the phrase "surface area." Indeed, a theorem about this result devised by Stephen Hawking is known as the *second law of black holes*:

In any natural process, the surface area of the event horizon of a black hole always increases or, at best, remains constant; it never decreases (in the absence or quantum effects).

The wording of this theorem deliberately imitates that of the second law of thermodynamics, which states that entropy always decreases, never increases. In fact, corresponding to the zeroth, first, second, and third law of thermodynamics (Box 4.1), theorems have been derived by Bardeen, Carter, Hawking, and Penrose called the zeroth, first, second, and third law of black-hole dynamics.

Problem 7.9. Let us use the second law of black-hole dynamics to calculate the maximum efficiency of black-hole processes as the second law of thermodynamics can be used to calculate the maximum efficiency of heat engines. As an example, consider the problem of the maximum energy extractable—say, in the form of gravitational radiation—from the merger of two black holes of equal mass M . Assume a head-on collision so that the initial and final black holes are nonrotating. In this simple situation, increase of the surface area implies

$$(16\pi G^2/c^4)M_{\text{final}}^2 > (16\pi G^2/c^4)M^2 + (16\pi G^2/c^4)M^2.$$

The energy released after all the commotion has died is given by $E = M_{\text{initial}}c^2 - M_{\text{final}}c^2$, where $M_{\text{initial}} = 2M$. The efficiency ε is defined to be $E/M_{\text{initial}}c^2$. From these considerations calculate the maximum possible value of ε . Numerical simulations show that the actual amount of gravitational radiation generated by head-on collisions of identical black holes involves on the order of 1 percent efficiency. How does this compare with the theoretical maximum?

The analogy between the second law of black-hole dynamics and the second law of thermodynamics led Jacob Bekenstein in 1972 to propose that the surface area of a black hole's event horizon is proportional to its entropy. This bold suggestion that a black hole is physically a *thermal* body was truly revolutionary, potentially as important for physics as the early thermodynamical discussions of the properties of blackbody radiation. (A black hole is the ultimate blackbody, since it absorbs every photon which falls on it.) For the idea to have merit, however, it must be possible: (a) to define precisely what is meant by the "entropy" of a black hole; and (b) to associate the concept of a temperature with a black hole.

Bekenstein proposed to define the entropy of the black hole on the basis of the theorem, "A black hole has no hair," which was proved by Carter, Hawking, Israel, and Robinson. This theorem has the following meaning. After a black hole is created by gravitational collapse and settles into an equilibrium configuration, an outside

