11

Vectors

11–1 Symmetry in physics

In this chapter we introduce a subject that is technically known in physics as *symmetry in physical law.* The word "symmetry" is used here with a special meaning, and therefore needs to be defined. When is a thing symmetrical—how can we define it? When we have a picture that is symmetrical, one side is somehow the same as the other side. Professor Hermann Weyl has given this definition of symmetry: a thing is symmetrical if one can subject it to a certain operation and it appears exactly the same after the operation. For instance, if we look at a vase that is left-and-right symmetrical, then turn it 180° around the vertical axis, it looks the same. We shall adopt the definition of symmetry in Weyl's more general form, and in that form we shall discuss symmetry of physical laws.

Suppose we build a complex machine in a certain place, with a lot of complicated interactions, and balls bouncing around with forces between them, and so on. Now suppose we build exactly the same kind of equipment at some other place, matching part by part, with the same dimensions and the same orientation, everything the same only displaced laterally by some distance. Then, if we start the two machines in the same initial circumstances, in exact correspondence, we ask: will one machine behave exactly the same as the other? Will it follow all the motions in exact parallelism? Of course the answer may well be *no,* because if we choose the wrong place for our machine it might be inside a wall and interferences from the wall would make the machine not work.

All of our ideas in physics require a certain amount of common sense in their application; they are not purely mathematical or abstract ideas. We have to understand what we mean when we say that the phenomena are the same when we move the apparatus to a new position. We mean that we move everything that we believe is relevant; if the phenomenon is not the same, we suggest that something relevant has not been moved, and we proceed to look for it. If we never find it, then we claim that the laws of physics do not have this symmetry. On the other hand, we may find it—we expect to find it—if the laws of physics do have this symmetry; looking around, we may discover, for instance, that the wall is pushing on the apparatus. The basic question is, if we define things well enough, if all the essential forces are included inside the apparatus, if all the relevant parts are moved from one place to another, will the laws be the same? Will the machinery work the same way?

It is clear that what we want to do is to move all the equipment and *essential* influences, but not *everything* in the world—planets, stars, and all—for if we do that, we have the same phenomenon again for the trivial reason that we are right back where we started. No, we cannot move *everything.* But it turns out in practice that with a certain amount of intelligence about what to move, the machinery will work. In other words, if we do not go inside a wall, if we know the origin of the outside forces, and arrange that those are moved too, then the machinery *will* work the same in one location as in another.

11–2 Translations

We shall limit our analysis to just mechanics, for which we now have sufficient knowledge. In previous chapters we have seen that the laws of mechanics can be summarized by a set of three equations for each particle:

$$m(d^2x/dt^2) = F_x, \qquad m(d^2y/dt^2) = F_y, \qquad m(d^2z/dt^2) = F_z. \tag{11.1}$$

Now this means that there exists a way to *measure* x, y, and z on three perpendicular axes, and the forces along those directions, such that these laws are true. These must be measured from some origin, but *where do we put the origin?* All that Newton would tell us at first is that there *is* some place that we can measure from, perhaps the center of the universe, such that these laws are correct. But we can show immediately that we can never find the center, because if we use some other origin it would make no difference. In other words, suppose that there are two people—Joe, who has an origin in one place, and Moe, who has a parallel system whose origin is somewhere else (Fig. 11–1). Now when Joe measures the location of the point in space, he finds it at x, y, and z (we shall usually leave z out because it is too confusing to draw in a picture). Moe, on the other hand, when measuring the same point, will obtain a different x (in order to distinguish it, we will call it x'), and in principle a different y, although in our example they are numerically equal. So we have

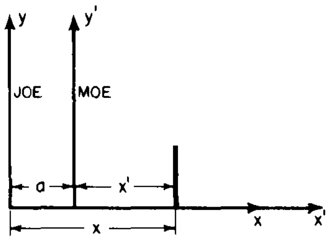

Fig. 11–1. Two parallel coordinate systems.

$$x' = x - a, \qquad y' = y, \qquad z' = z. \tag{11.2}$$

Now in order to complete our analysis we must know what Moe would obtain for the forces. The force is supposed to act along some line, and by the force in the x-direction we mean the part of the total which is in the x-direction, which is the magnitude of the force times this cosine of its angle with the x-axis. Now we see that Moe would use exactly the same projection as Joe would use, so we have a set of equations

$$F_{x'} = F_x, \qquad F_{y'} = F_y, \qquad F_{z'} = F_z. \tag{11.3}$$

These would be the relationships between quantities as seen by Joe and Moe.

The question is, if Joe knows Newton's laws, and if Moe tries to write down Newton's laws, will they also be correct for him? Does it make any difference from which origin we measure the points? In other words, assuming that equations (11.1) are true, and the Eqs. (11.2) and (11.3) give the relationship of the measurements, is it or is it not true that

$$\begin{aligned}
&\text{(a)} \quad m(d^2x'/dt^2) = F_{x'},\\
&\text{(b)} \quad m(d^2y'/dt^2) = F_{y'},\\
&\text{(c)} \quad m(d^2z'/dt^2) = F_{z'}?
\end{aligned} \tag{11.4}$$

In order to test these equations we shall differentiate the formula for x' twice. First of all

$$\frac{dx'}{dt} = \frac{d}{dt}(x - a) = \frac{dx}{dt} - \frac{da}{dt}.$$

Now we shall assume that Moe's origin is fixed (not moving) relative to Joe's; therefore a is a constant and $da/dt = 0$, so we find that

$$dx'/dt = dx/dt'$$

and therefore

$$d^2x'/dt^2 = d^2x/dt^2;$$

therefore we know that Eq. (11.4a) becomes

$$m(d^2x/dt^2) = F_{x'}.$$

(We also suppose that the masses measured by Joe and Moe are equal.) Thus the acceleration times the mass is the same as the other fellow's. We have also found the formula for $F_{x'}$, for, substituting from Eq. (11.1), we find that

$$F_{x'} = F_x.$$

Therefore the laws as seen by Moe appear the same; he can write Newton's laws too, with different coordinates, and they will still be right. That means that there is no unique way to define the origin of the world, because the laws will appear the same, from whatever position they are observed.

This is also true: if there is a piece of equipment in one place with a certain kind of machinery in it, the same equipment in another place will behave in the same way. Why? Because one machine, when analyzed by Moe, has exactly the same equations as the other one, analyzed by Joe. Since the *equations* are the same, the *phenomena* appear the same. So the proof that an apparatus in a new position behaves the same as it did in the old position is the same as the proof that the equations when displaced in space reproduce themselves. Therefore we say that *the laws of physics are symmetrical for translational displacements,* symmetrical in the sense that the laws do not change when we make a translation of our coordinates. Of course it is quite obvious intuitively that this is true, but it is interesting and entertaining to discuss the mathematics of it.

11-3 Rotations

The above is the first of a series of ever more complicated propositions concerning the symmetry of a physical law. The next proposition is that it should make no difference in which *direction* we choose the axes. In other words, if we build a piece of equipment in some place and watch it operate, and nearby we build the same kind of apparatus but put it up on an angle, will it operate in the same way? Obviously it will not if it is a Grandfather clock, for example! If a pendulum clock stands upright, it works fine, but if it is tilted the pendulum falls against the side of the case and nothing happens. The theorem is then false in the case of the pendulum clock, unless we include the earth, which is pulling on the pendulum. Therefore we can make a prediction about pendulum clocks if we believe in the symmetry of physical law for rotation: something else is involved in the operation of a pendulum clock besides the machinery of the clock, something outside it that we should look for. We may also predict that pendulum clocks will not work the same way when located in different places relative to this mysterious source of asymmetry, perhaps the earth. Indeed, we know that a pendulum clock up in an artificial satellite, for example, would not tick either, because there is no effective force, and on Mars it would go at a different rate. Pendulum clocks *do* involve something more than just the machinery inside, they involve something on the outside. Once we recognize this factor, we see that we must turn the earth along with the apparatus. Of course we do not have to worry about that, it is easy to do; one simply waits a moment or two and the earth turns; then the pendulum clock ticks again in the new position the same as it did before. While we are rotating in space our angles are always changing, absolutely; this change does not seem to bother us very much, for in the new position we seem to be in the same condition as in the old. This has a certain tendency to confuse one, because it is true that in the new turned position the laws are the same as in the unturned position, but it is *not* true that *as we turn* a thing it follows the same laws as it does when we are not turning it. If we perform sufficiently delicate experiments, we can tell that the earth *is rotating,* but not that it *had rotated.* In other words, we cannot locate its angular position, but we can tell that it is changing.

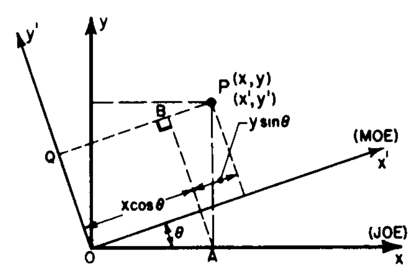

Fig. 11–2. Two coordinate systems having different angular orientations.

Now we may discuss the effects of angular orientation upon physical laws. Let us find out whether the same game with Joe and Moe works again. This time, to avoid needless complication, we shall suppose that Joe and Moe use the same origin (we have already shown that the axes can be moved by translation to another place). Assume that Moe's axes have rotated relative to Joe's by an angle θ. The two coordinate systems are shown in Fig. 11–2, which is restricted to two dimensions. Consider any point P having coordinates (x, y) in Joe's system and (x', y') in Moe's system. We shall begin, as in the previous case, by expressing the coordinates x' and y' in terms of x, y, and θ. To do so, we first drop perpendiculars from P to all four axes and draw AB perpendicular to PQ. Inspection of the figure shows that x' can be written as the sum of two lengths along the x'-axis, and y' as the difference of two lengths along AB. All these lengths are expressed in terms of x, y, and θ in equations (11.5), to which we have added an equation for the third dimension.

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta, \\ y' &= y \cos \theta - x \sin \theta, \\ z' &= z. \end{aligned} \tag{11.5}$$

The next step is to analyze the relationship of forces as seen by the two observers, following the same general method as before. Let us assume that a force F, which has already been analyzed as having components F_x and F_y (as seen by Joe), is acting on a particle of mass m, located at point P in Fig. 11–2. For simplicity, let us move both sets of axes so that the origin is at P, as shown in Fig. 11–3. Moe sees the components of F along his axes as $F_{x'}$ and $F_{y'}$. F_x has components along both the x'- and y'-axes, and F_y likewise has components along both these axes. To express $F_{x'}$ in terms of F_x and F_y, we sum these components along the x'-axis, and in a like manner we can express $F_{y'}$ in terms of F_x and F_y. The results are

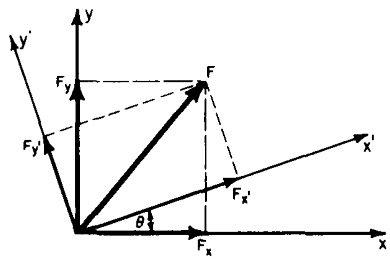

Fig. 11–3. Components of a force in the two systems.

$$\begin{aligned} F_{x'} &= F_x \cos \theta + F_y \sin \theta, \\ F_{y'} &= F_y \cos \theta - F_x \sin \theta, \\ F_{z'} &= F_z. \end{aligned} \tag{11.6}$$

It is interesting to note an accident of sorts, which is of extreme importance: the formulas (11.5) and (11.6), for coordinates of P and components of F, respectively, *are of identical form.*

As before, Newton's laws are assumed to be true in Joe's system, and are expressed by equations (11.1). The question, again, is whether Moe can apply Newton's laws—will the results be correct for his system of rotated axes? In other words, if we assume that Eqs. (11.5) and (11.6) give the relationship of the measurements, is it true or not true that

$$\begin{aligned} m(d^2x'/dt^2) &= F_{x'}, \\ m(d^2y'/dt^2) &= F_{y'}, \\ m(d^2z'/dt^2) &= F_{z'}? \end{aligned} \tag{11.7}$$

To test these equations, we calculate the left and right sides independently, and compare the results. To calculate the left sides, we multiply equations (11.5) by m, and differentiate twice with respect to time, assuming the angle θ to be constant. This gives

$$\begin{aligned} m(d^2x'/dt^2) &= m(d^2x/dt^2) \cos \theta + m(d^2y/dt^2) \sin \theta, \\ m(d^2y'/dt^2) &= m(d^2y/dt^2) \cos \theta - m(d^2x/dt^2) \sin \theta, \\ m(d^2z'/dt^2) &= m(d^2z/dt^2). \end{aligned} \tag{11.8}$$

We calculate the right sides of equations (11.7) by substituting equations (11.1) into equations (11.6). This gives

$$\begin{aligned} F_{x'} &= m(d^2x/dt^2) \cos \theta + m(d^2y/dt^2) \sin \theta, \\ F_{y'} &= m(d^2y/dt^2) \cos \theta - m(d^2x/dt^2) \sin \theta, \\ F_{z'} &= m(d^2z/dt^2). \end{aligned} \tag{11.9}$$

Behold! The right sides of Eqs. (11.8) and (11.9) are identical, so we conclude that if Newton's laws are correct on one set of axes, they are also valid on any other set of axes. This result, which has now been established for both translation and rotation of axes, has certain consequences: first, no one can claim his particular axes are unique, but of course they can be more *convenient* for certain particular problems. For example, it is handy to have gravity along one axis, but this is not physically necessary. Second, it means that any piece of equipment which is completely self-contained, with all the force-generating equipment completely inside the apparatus, would work the same when turned at an angle.

11–4 Vectors

Not only Newton's laws, but also the other laws of physics, so far as we know today, have the two properties which we call invariance (or symmetry) under translation of axes and rotation of axes. These properties are so important that a mathematical technique has been developed to take advantage of them in writing and using physical laws.

The foregoing analysis involved considerable tedious mathematical work. To reduce the details to a minimum in the analysis of such questions, a very powerful mathematical machinery has been devised. This system, called *vector analysis*, supplies the title of this chapter; strictly speaking, however, this is a chapter on the symmetry of physical laws. By the methods of the preceding analysis we were able to do everything required for obtaining the results that we sought, but in practice we should like to do things more easily and rapidly, so we employ the vector technique.

We began by noting some characteristics of two kinds of quantities that are important in physics. (Actually there are more than two, but let us start out with two.) One of them, like the number of potatoes in a sack, we call an ordinary quantity, or an undirected quantity, or a *scalar*. Temperature is an example of such a quantity. Other quantities that are important in physics do have direction, for instance velocity: we have to keep track of which way a body is going, not just its speed. Momentum and force also have direction, as does displacement: when someone steps from one place to another in space, we can keep track of how far he went, but if we wish also to know *where* he went, we have to specify a direction.

All quantities that have a direction, like a step in space, are called *vectors*.

A vector is three numbers. In order to represent a step in space, say from the origin to some particular point P whose location is (x, y, z), we really need three numbers, but we are going to invent a single mathematical symbol, $\mathbf{r}$, which is unlike any other mathematical symbols we have so far used.* It is *not* a single number, it represents *three* numbers: x, y, and z. It means three numbers, but not really only *those* three numbers, because if we were to use a different coordinate system, the three numbers would be changed to x', y', and z'. However, we want to keep our mathematics simple and so we are going to use the *same mark* to represent the three numbers (x, y, z) and the three numbers (x', y', z'). That is, we use the same mark to represent the first set of three numbers for one coordinate system, but the second set of three numbers if we are using the other coordinate system. This has the advantage that when we change the coordinate system, we do not have to change the letters of our equations. If we write an equation in terms of x, y, z, and then use another system, we have to change to x', y', z', but we shall just write $\mathbf{r}$, with the convention that it represents (x, y, z) if we use one set of axes, or (x', y', z') if we use another set of axes, and so on. The three numbers which describe the quantity in a given coordinate system are called the *components* of the vector in the direction of the coordinate axes of that system. That is, we use the same symbol for the three letters that correspond to the *same object, as seen from different axes*. The very fact that we can say "the same object" implies a physical intuition about the reality of a step in space, that is independent of the components in terms of which we measure it. So the symbol $\mathbf{r}$ will represent the same thing no matter how we turn the axes.

Now suppose there is another directed physical quantity, any other quantity, which also has three numbers associated with it, like force, and these three numbers change to three other numbers by a certain mathematical rule, if we change the axes. It must be the same rule that changes (x, y, z) into (x', y', z'). In other words, any physical quantity associated with three numbers which transform as do the components of a step in space is a vector. An equation like

$$\mathbf{F} = \mathbf{r}$$

would thus be true in *any* coordinate system if it were true in one. This equation,

---

* In type, vectors are represented by boldface; in handwritten form an arrow is used: $\vec{r}$.

of course, stands for the three equations

$$F_x = x, \qquad F_y = y, \qquad F_z = z,$$

or, alternatively, for

$$F_{x'} = x', \qquad F_{y'} = y', \qquad F_{z'} = z'.$$

The fact that a physical relationship can be expressed as a vector equation assures us the relationship is unchanged by a mere rotation of the coordinate system. That is the reason why vectors are so useful in physics.

Now let us examine some of the properties of vectors. As examples of vectors we may mention velocity, momentum, force, and acceleration. For many purposes it is convenient to represent a vector quantity by an arrow that indicates the direction in which it is acting. Why can we represent force, say, by an arrow? Because it has the same mathematical transformation properties as a "step in space." We thus represent it in a diagram as if it were a step, using a scale such that one unit of force, or one newton, corresponds to a certain convenient length. Once we have done this, all forces can be represented as lengths, because an equation like

$$\mathbf{F} = k\mathbf{r},$$

where k is some constant, is a perfectly legitimate equation. Thus we can always represent forces by lines, which is very convenient, because once we have drawn the line we no longer need the axes. Of course, we can quickly calculate the three components as they change upon turning the axes, because that is just a geometric problem.

11–5 Vector algebra

Now we must describe the laws, or rules, for combining vectors in various ways. The first such combination is the *addition* of two vectors: suppose that $\mathbf{a}$ is a vector which in some particular coordinate system has the three components (a_x, a_y, a_z), and that $\mathbf{b}$ is another vector which has the three components (b_x, b_y, b_z). Now let us invent three new numbers $(a_x + b_x, a_y + b_y, a_z + b_z)$. Do these form a vector? "Well," we might say, "they are three numbers, and every three numbers form a vector." No, *not* every three numbers form a vector! In order for it to be a vector, not only must there be three numbers, but these must be associated with a coordinate system in such a way that if we turn the coordinate system, the three numbers "revolve" on each other, get "mixed up" in each other, by the precise laws we have already described. So the question is, if we now rotate the coordinate system so that (a_x, a_y, a_z) become $(a_{x'}, a_{y'}, a_{z'})$ and (b_x, b_y, b_z) become $(b_{x'}, b_{y'}, b_{z'})$, what do $(a_x + b_x, a_y + b_y, a_z + b_z)$ become? Do they become $(a_{x'} + b_{x'}, a_{y'} + b_{y'}, a_{z'} + b_{z'})$ or not? The answer is, of course, yes, because the prototype transformations of Eq. (11.5) constitute what we call a *linear* transformation. If we apply those transformations to a_x and b_x to get $a_{x'} + b_{x'}$, we find that the transformed $a_x + b_x$ is indeed the same as $a_{x'} + b_{x'}$. When $\mathbf{a}$ and $\mathbf{b}$ are "added together" in this sense, they will form a vector which we may call $\mathbf{c}$. We would write this as

$$\mathbf{c} = \mathbf{a} + \mathbf{b}.$$

Now $\mathbf{c}$ has the interesting property

$$\mathbf{c} = \mathbf{b} + \mathbf{a},$$

as we can immediately see from its components. Thus also,

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}.$$

We can add vectors in any order.

What is the geometric significance of $\mathbf{a} + \mathbf{b}$? Suppose that $\mathbf{a}$ and $\mathbf{b}$ were represented by lines on a piece of paper, what would $\mathbf{c}$ look like? This is shown in

Fig. 11–4. We see that we can add the components of **b** to those of **a** most conveniently if we place the rectangle representing the components of **b** next to that representing the components of **a** in the manner indicated. Since **b** just "fits" into its rectangle, as does **a** into its rectangle, this is the same as putting the "tail" of **b** on the "head" of **a**, the arrow from the "tail" of **a** to the "head" of **b** being the vector **c**. Of course, if we added **a** to **b** the other way around, we would put the "tail" of **a** on the "head" of **b**, and by the geometrical properties of parallelograms we would get the same result for **c**. Note that vectors can be added in this way without reference to any coordinate axes.

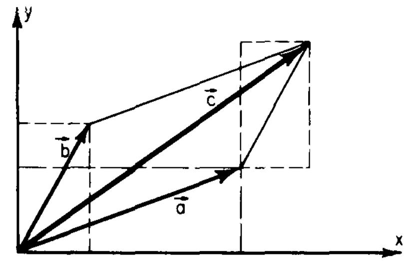

Fig. 11–4. The addition of vectors.

Suppose we multiply a vector by a number α, what does this mean? We *define* it to mean a new vector whose components are αa_x, αa_y, and αa_z. We leave it as a problem for the student to prove that it *is* a vector.

Now let us consider vector subtraction. We may define subtraction in the same way as addition, but instead of adding, we subtract the components. Or we might define subtraction by defining a negative vector, $-\mathbf{b} = -1\mathbf{b}$, and then we would add the components. It comes to the same thing. The result is shown in Fig. 11–5. This figure shows $\mathbf{d} = \mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$; we also note that the difference $\mathbf{a} - \mathbf{b}$ can be found very easily from **a** and **b** by using the equivalent relation $\mathbf{a} = \mathbf{b} + \mathbf{d}$. Thus the difference is even easier to find than the sum: we just draw the vector from **b** to **a**, to get $\mathbf{a} - \mathbf{b}$!

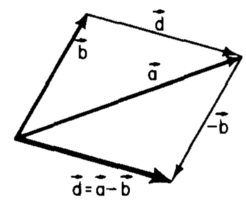

Fig. 11–5. The subtraction of vectors.

Next we discuss velocity. Why is velocity a vector? If position is given by the three coordinates (x, y, z), what is the velocity? The velocity is given by dx/dt, dy/dt, and dz/dt Is that a vector, or not? We can find out by differentiating the expressions in Eq. (11.5) to find out whether dx'/dt *transforms* in the right way. We see that the components dx/dt and dy/dt *do* transform according to the same law as x and y, and therefore the time derivative is a vector. So the velocity *is* a vector. We can write the velocity in an interesting way as

$$\mathbf{v} = d\mathbf{r}/dt.$$

What the velocity is, and why it is a vector, can also be understood more pictorially: How far does a particle move in a short time Δt? Answer: $\Delta\mathbf{r}$, so if a particle is "here" at one instant and "there" at another instant, then the vector difference of the positions $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, which is in the direction of motion shown in Fig. 11–6, divided by the time interval $\Delta t = t_2 - t_1$, is the "average velocity" vector.

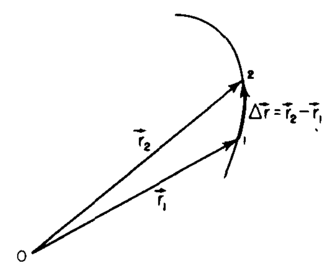

Fig. 11–6. The displacement of a particle in a short time interval $\Delta t = t_2 - t_1$.

In other words, by vector velocity we mean the limit, as Δt goes to 0, of the difference between the radius vectors at the time $t + \Delta t$ and the time t, divided by Δt:

$$\mathbf{v} = \lim_{\Delta t \to 0} (\Delta\mathbf{r}/\Delta t) = d\mathbf{r}/dt. \tag{11.10}$$

Thus velocity is a vector because it is the difference of two vectors. It is also the right definition of velocity because its components are dx/dt, dy/dt, and dz/dt. In fact, we see from this argument that if we differentiate *any* vector with respect to time we produce a new vector. So we have several ways of producing new vectors: (1) multiply by a constant, (2) differentiate with respect to time, (3) add or subtract two vectors.

11–6 Newton's laws in vector notation

In order to write Newton's laws in vector form, we have to go just one step further, and define the acceleration vector. This is the time derivative of the velocity vector, and it is easy to demonstrate that its components are the second derivatives of x, y, and z with respect to t:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \left(\frac{d}{dt}\right)\left(\frac{d\mathbf{r}}{dt}\right) = \frac{d^2\mathbf{r}}{dt^2}, \tag{11.11}$$

$$a_x = \frac{dv_x}{dt} = \frac{d^2x}{dt^2}, \quad a_y = \frac{dv_y}{dt} = \frac{d^2y}{dt^2}, \quad a_z = \frac{dv_z}{dt} = \frac{d^2z}{dt^2}. \tag{11.12}$$

With this definition, then, Newton's laws can be written in this way:

$$m\mathbf{a} = \mathbf{F} \tag{11.13}$$

or

$$m(d^2\mathbf{r}/dt^2) = \mathbf{F}. \tag{11.14}$$

Now the problem of proving the invariance of Newton's laws under rotation of coordinates is this: prove that **a** is a vector; this we have just done. Prove that **F** is a vector; we *suppose* it is. So if force is a vector, then, since we know acceleration is a vector, Eq. (11.13) will look the same in any coordinate system. Writing it in a form which does not explicitly contain x's, y's, and z's has the advantage that from now on we need not write *three* laws every time we write Newton's equations or other laws of physics. We write what looks like *one* law, but really, of course, it is the three laws for any particular set of axes, because any vector equation involves the statement that *each of the components is equal.*

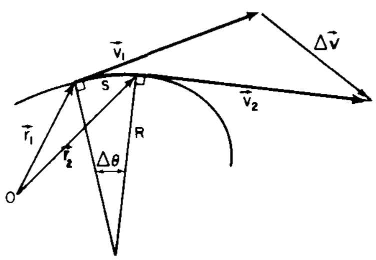

Fig. 11–7. A curved trajectory.

The fact that the acceleration is the rate of change of the vector velocity helps us to calculate the acceleration in some rather complicated circumstances. Suppose, for instance, that a particle is moving on some complicated curve (Fig. 11–7) and that, at a given instant t, it had a certain velocity $\mathbf{v}_1$, but that when we go to another instant t_2 a little later, it has a different velocity $\mathbf{v}_2$. What is the acceleration? Answer: Acceleration is the difference in the velocity divided by the small time interval, so we need the difference of the two velocities. How do we get the difference of the velocities? To subtract two vectors, we put the vector across the ends of $\mathbf{v}_2$ and $\mathbf{v}_1$; that is, we draw $\mathbf{\Delta}$ as the difference of the two vectors, right? *No!* That only works when the *tails* of the vectors are in the same place! It has no meaning if we move the vector somewhere else and then draw a line across, so watch out! We have to draw a new diagram to subtract the vectors. In Fig. 11–8, $\mathbf{v}_1$ and $\mathbf{v}_2$ are both drawn parallel and equal to their counterparts in Fig. 11–7, and now we can discuss the acceleration. Of course the acceleration is simply $\Delta\mathbf{v}/\Delta t$. It is interesting to note that we can compose the velocity difference out of two parts; we can think of acceleration as having *two components*, $\Delta\mathbf{v}_{\parallel}$ in the direction tangent to the path and $\Delta\mathbf{v}_{\perp}$ at right angles to the path, as indicated in Fig. 11–8. The acceleration tangent to the path is, of course, just the change in the *length* of the vector, i.e., the change in the *speed* v:

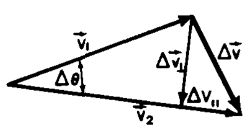

Fig. 11–8. Diagram for calculating the acceleration.

$$a_{\parallel} = dv/dt. \tag{11.15}$$

The other component of acceleration, at right angles to the curve, is easy to calculate, using Figs. 11–7 and 11–8. In the short time Δt let the change in angle between $\mathbf{v}_1$ and $\mathbf{v}_2$ be the small angle $\Delta\theta$. If the magnitude of the velocity is called v, then of course

$$\Delta v_{\perp} = v\,\Delta\theta$$

and the acceleration a will be

$$a_{\perp} = v\,(\Delta\theta/\Delta t).$$

Now we need to know $\Delta\theta/\Delta t$, which can be found this way: If, at the given moment, the curve is approximated as a circle of a certain radius R, then in a time Δt the distance s is, of course, $v\,\Delta t$, where v is the speed.

$$\Delta\theta = v\,(\Delta t/R), \qquad \text{or} \qquad \Delta\theta/\Delta t = v/R.$$

Therefore, we find

$$a = v^2/R, \tag{11.16}$$

as we have seen before.

11–7 Scalar product of vectors

Now let us examine a little further the properties of vectors. It is easy to see that the *length* of a step in space would be the same in any coordinate system. That is, if a particular step $\mathbf{r}$ is represented by x, y, z, in one coordinate system, and by x', y', z' in another coordinate system, surely the distance $r = |\mathbf{r}|$ would be the same in both. Now

$$r = \sqrt{x^2 + y^2 + z^2}$$

and also

$$r' = \sqrt{x'^2 + y'^2 + z'^2}.$$

So what we wish to verify is that these two quantities are equal. It is much more convenient not to bother to take the square root, so let us talk about the square of the distance; that is, let us find out whether

$$x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2. \tag{11.17}$$

It had better be—and if we substitute Eq. (11.5) we do indeed find that it is. So we see that there are other kinds of equations which are true for any two coordinate systems.

Something new is involved. We can produce a new quantity, a function of x, y, and z, called a *scalar function*, a quantity which has no direction but which is the same in both systems. Out of a vector we can make a scalar. We have to find a general rule for that. It is clear what the rule is for the case just considered: add the squares of the components. Let us now define a new thing, which we call $\mathbf{a} \cdot \mathbf{a}$. This is not a vector, but a scalar; it is a number that is the same in all coordinate systems, and it is defined to be the sum of the squares of the three components of the vector:

$$\mathbf{a} \cdot \mathbf{a} = a_x^2 + a_y^2 + a_z^2. \tag{11.18}$$

Now you say, "But with what axes?" It does not depend on the axes, the answer is the same in *every* set of axes. So we have a new *kind* of quantity, a new *invariant* or *scalar* produced by one vector "squared." If we now define the following quantity for any two vectors $\mathbf{a}$ and $\mathbf{b}$:

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z, \tag{11.19}$$

we find that this quantity, calculated in the primed and unprimed systems, also stays the same. To prove it we note that it is true of $\mathbf{a} \cdot \mathbf{a}$, $\mathbf{b} \cdot \mathbf{b}$, and $\mathbf{c} \cdot \mathbf{c}$, where $\mathbf{c} = \mathbf{a} + \mathbf{b}$. Therefore the sum of the squares $(a_x + b_x)^2 + (a_y + b_y)^2 + (a_z + b_z)^2$ will be invariant:

$$\begin{aligned}(a_x + b_x)^2 + (a_y + b_y)^2 + (a_z + b_z)^2 = (a_{x'} + b_{x'})^2 \\ + (a_{y'} + b_{y'})^2 + (a_{z'} + b_{z'})^2.\end{aligned} \tag{11.20}$$

If both sides of this equation are expanded, there will be cross products of just the type appearing in Eq. (11.19), as well as the sums of squares of the components of $\mathbf{a}$ and $\mathbf{b}$. The invariance of terms of the form of Eq. (11.18) then leaves the cross product terms (11.19) invariant also.

The quantity $\mathbf{a} \cdot \mathbf{b}$ is called the *scalar product* of two vectors, $\mathbf{a}$ and $\mathbf{b}$, and it has many interesting and useful properties. For instance, it is easily proved that

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}. \tag{11.21}$$

Also, there is a simple geometrical way to calculate $\mathbf{a} \cdot \mathbf{b}$, without having to calculate the components of $\mathbf{a}$ and $\mathbf{b}$: $\mathbf{a} \cdot \mathbf{b}$ is the product of the length of $\mathbf{a}$ and the length of $\mathbf{b}$ times the cosine of the angle between them. Why? Suppose that we choose a special coordinate system in which the x-axis lies along $\mathbf{a}$; in those circumstances, the only component of $\mathbf{a}$ that will be there is a_x, which is of course the whole length of $\mathbf{a}$. Thus Eq. (11.19) reduces to $\mathbf{a} \cdot \mathbf{b} = a_x b_x$ for this case, and this is the length of $\mathbf{a}$ times the component of $\mathbf{b}$ in the direction of $\mathbf{a}$, that is, $b \cos \theta$:

$$\mathbf{a} \cdot \mathbf{b} = ab \cos \theta.$$

Therefore, in that special coordinate system, we have proved that $\mathbf{a} \cdot \mathbf{b}$ is the length of **a** times the length of **b** times cos θ. But *if it is true in one coordinate system, it is true in all*, because $\mathbf{a} \cdot \mathbf{b}$ is independent of the coordinate system; that is our argument.

What good is the dot product? Are there any cases in physics where we need it? Yes, we need it all the time. For instance, in Chapter 4 the kinetic energy was called $\frac{1}{2}mv^2$, but if the object is moving in space it should be the velocity squared in the x-direction, the y-direction, and the z-direction, and so the formula for kinetic energy according to vector analysis is

$$\text{K.E.} = \tfrac{1}{2}m(\mathbf{v} \cdot \mathbf{v}) = \tfrac{1}{2}m(v_x^2 + v_y^2 + v_z^2). \tag{11.22}$$

Energy does not have direction. Momentum has direction; it is a vector, and it is the mass times the velocity vector.

Another example of a dot product is the work done by a force when something is pushed from one place to the other. We have not yet defined work, but it is equivalent to the energy change, the weights lifted, when a force **F** acts through a distance **s**:

$$\text{Work} = \mathbf{F} \cdot \mathbf{s} \tag{11.23}$$

It is sometimes very convenient to talk about the component of a vector in a certain direction (say the vertical direction because that is the direction of gravity). For such purposes, it is useful to invent what we call a *unit vector* in the direction that we want to study. By a unit vector we mean one whose dot product with itself is equal to unity. Let us call this unit vector **i**; then $\mathbf{i} \cdot \mathbf{i} = 1$. Then, if we want the component of some vector in the direction of **i**, we see that the dot product $\mathbf{a} \cdot \mathbf{i}$ will be $a \cos \theta$, i.e., the component of **a** in the direction of **i**. This is a nice way to get the component; in fact, it permits us to get *all* the components and to write a rather amusing formula. Suppose that in a given system of coordinates, x, y, and z, we invent three vectors: **i**, a unit vector in the direction x; **j**, a unit vector in the direction y; and **k**, a unit vector in the direction z. Note first that $\mathbf{i} \cdot \mathbf{i} = 1$. What is $\mathbf{i} \cdot \mathbf{j}$? When two vectors are at right angles, their dot product is zero. Thus

$$\begin{array}{lll} \mathbf{i} \cdot \mathbf{i} = 1 & & \\ \mathbf{i} \cdot \mathbf{j} = 0 & \mathbf{j} \cdot \mathbf{j} = 1 & \\ \mathbf{i} \cdot \mathbf{k} = 0 & \mathbf{j} \cdot \mathbf{k} = 0 & \mathbf{k} \cdot \mathbf{k} = 1 \end{array} \tag{11.24}$$

Now with these definitions, any vector whatsoever can be written this way:

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}. \tag{11.25}$$

By this means we can go from the components of a vector to the vector itself.

This discussion of vectors is by no means complete. However, rather than try to go more deeply into the subject now, we shall first learn to use in physical situations some of the ideas so far discussed. Then, when we have properly mastered this basic material, we shall find it easier to penetrate more deeply into the subject without getting too confused. We shall later find that it is useful to define another kind of product of two vectors, called the vector product, and written as $\mathbf{a} \times \mathbf{b}$. However, we shall undertake a discussion of such matters in a later chapter.

12

Characteristics of Force

12–1 What is a force?

Although it is interesting and worth while to study the physical laws simply because they help us to understand and to use nature, one ought to stop every once in a while and think, "What do they really mean?" The meaning of any statement is a subject that has interested and troubled philosophers from time immemorial, and the meaning of physical laws is even more interesting, because it is generally believed that these laws represent some kind of real knowledge. The meaning of knowledge is a deep problem in philosophy, and it is always important to ask, "What does it mean?"

Let us ask, "What is the meaning of the physical laws of Newton, which we write as $F = ma$? What is the meaning of force, mass, and acceleration?" Well, we can intuitively sense the meaning of mass, and we can *define* acceleration if we know the meaning of position and time. We shall not discuss those meanings, but shall concentrate on the new concept of *force*. The answer is equally simple: "If a body is accelerating, then there is a force on it." That is what Newton's laws say, so the most precise and beautiful definition of force imaginable might simply be to say that force is the mass of an object times the acceleration. Suppose we have a law which says that the conservation of momentum is valid if the sum of all the external forces is zero; then the question arises, "What does it *mean*, that the sum of all the external forces is zero?" A pleasant way to define that statement would be: "When the total momentum is a constant, then the sum of the external forces is zero." There must be something wrong with that, because it is just not saying anything new. If we have discovered a fundamental law, which asserts that the force is equal to the mass times the acceleration, and then *define* the force to be the mass times the acceleration, we have found out nothing. We could also define force to mean that a moving object with no force acting on it continues to move with constant velocity in a straight line. If we then observe an object *not* moving in a straight line with a constant velocity, we might say that there is a force on it. Now such things certainly cannot be the content of physics, because they are definitions going in a circle. The Newtonian statement above, however, seems to be a most precise definition of force, and one that appeals to the mathematician; nevertheless, it is completely useless, because no prediction whatsoever can be made from a definition. One might sit in an armchair all day long and define words at will, but to find out what happens when two balls push against each other, or when a weight is hung on a spring, is another matter altogether, because the way the bodies *behave* is something completely outside any choice of definitions.

For example, if we were to choose to say that an object left to itself keeps its position and does not move, then when we see something drifting, we could say that must be due to a "gorce"—a gorce is the rate of change of position. Now we have a wonderful new law, everything stands still except when a gorce is acting. You see, that would be analogous to the above definition of force, and it would contain no information. The real content of Newton's laws is this: that the force is supposed to have some *independent properties*, in addition to the law $F = ma$; but the *specific* independent properties that the force has were not completely described by Newton or by anybody else, and therefore the physical law $F = ma$ is an incomplete law. It implies that if we study the mass times the acceleration and call the product the force, i.e., if we study the characteristics of force as a program of interest, then we shall find that forces have some simplicity; the law is a good program for analyzing nature, it is a suggestion that the forces will be simple.

Now the first example of such forces was the complete law of gravitation, which was given by Newton, and in stating the law he answered the question, "What is the force?" If there were nothing but gravitation, then the combination of this law and the force law (second law of motion) would be a complete theory, but there is much more than gravitation, and we want to use Newton's laws in many different situations. Therefore in order to proceed we have to tell something about the properties of force.

For example, in dealing with force the tacit assumption is always made that the force is equal to zero unless some physical body is present, that if we find a force that is not equal to zero we also find something in the neighborhood that is a source of the force. This assumption is entirely different from the case of the "gorce" that we introduced above. One of the most important characteristics of force is that it has a material origin, and this is *not* just a definition.

Newton also gave one rule about the force: that the forces between interacting bodies are equal and opposite—action equals reaction; that rule, it turns out, is not exactly true. In fact, the law $F = ma$ is not exactly true; if it were a definition we should have to say that it is *always* exactly true; but it is not.

The student may object, "I do not like this imprecision, I should like to have everything defined exactly; in fact, it says in some books that any science is an exact subject, in which *everything* is defined." If you insist upon a precise definition of force, you will never get it! First, because Newton's Second Law is not exact, and second, because in order to understand physical laws you must understand that they are all some kind of approximation.

Any simple idea is approximate; as an illustration, consider an object, ... what *is* an object? Philosophers are always saying, "Well, just take a chair for example." The moment they say that, you know that they do not know what they are talking about any more. What *is* a chair? Well, a chair is a certain thing over there ... certain?, how certain? The atoms are evaporating from it from time to time—not many atoms, but a few—dirt falls on it and gets dissolved in the paint; so to define a chair precisely, to say exactly which atoms are chair, and which atoms are air, or which atoms are dirt, or which atoms are paint that belongs to the chair is impossible. So the mass of a chair can be defined only approximately. In the same way, to define the mass of a single object is impossible, because there are not any single, left-alone objects in the world—every object is a mixture of a lot of things, so we can deal with it only as a series of approximations and idealizations.

The trick is the idealizations. To an excellent approximation of perhaps one part in 10^{10}, the number of atoms in the chair does not change in a minute, and if we are not too precise we may idealize the chair as a definite thing; in the same way we shall learn about the characteristics of force, in an ideal fashion, if we are not too precise. One may be dissatisfied with the approximate view of nature that physics tries to obtain (the attempt is always to increase the accuracy of the approximation), and may prefer a mathematical definition; but mathematical definitions can never work in the real world. A mathematical definition will be good for mathematics, in which all the logic can be followed out completely, but the physical world is complex, as we have indicated in a number of examples, such as those of the ocean waves and a glass of wine. When we try to isolate pieces of it, to talk about one mass, the wine and the glass, how can we know which is which, when one dissolves in the other? The forces on a single thing already involve approximation, and if we have a system of discourse about the real world, then that system, at least for the present day, must involve approximations of some kind.

This system is quite unlike the case of mathematics, in which everything can be defined, and then we do not *know* what we are talking about. In fact, the glory of mathematics is that *we do not have to say what we are talking about.* The glory is that the laws, the arguments, and the logic are independent of what "it" is. If we have any other set of objects that obey the same system of axioms as Euclid's geometry, then if we make new definitions and follow them out with correct logic, all the consequences will be correct, and it makes no difference what the subject was. In nature, however, when we draw a line or establish a line by using a light beam and a theodolite, as we do in surveying, are we measuring a line in the sense of Euclid? No, we are making an approximation; the cross hair has some width, but a geometrical line has no width, and so, whether Euclidean geometry can be used for surveying or not is a physical question, not a mathematical question. However, from an experimental standpoint, not a mathematical standpoint, we need to know whether the laws of Euclid apply to the kind of geometry that we use in measuring land; so we make a hypothesis that it does, and it works pretty well; but it is not precise, because our surveying lines are not really geometrical lines. Whether or not those lines of Euclid, which are really abstract, apply to the lines of experience is a question for experience; it is not a question that can be answered by sheer reason.

In the same way, we cannot just call $F = ma$ a definition, deduce everything purely mathematically, and make mechanics a mathematical theory, when mechanics is a description of nature. By establishing suitable postulates it is always possible to make a system of mathematics, just as Euclid did, but we cannot make a mathematics of the world, because sooner or later we have to find out whether the axioms are valid for the objects of nature. Thus we immediately get involved with these complicated and "dirty" objects of nature, but with approximations ever increasing in accuracy.

12–2 Friction

The foregoing considerations show that a true understanding of Newton's laws requires a discussion of forces, and it is the purpose of this chapter to introduce such a discussion, as a kind of completion of Newton's laws. We have already studied the definitions of acceleration and related ideas, but now we have to study the properties of force, and this chapter, unlike the previous chapters, will not be very precise, because forces are quite complicated.

To begin with a particular force, let us consider the drag on an airplane flying through the air. What is the law for that force? (Surely there is a law for every force, we *must* have a law!) One can hardly think that the law for that force will be simple. Try to imagine what makes a drag on an airplane flying through the air—the air rushing over the wings, the swirling in the back, the changes going on around the fuselage, and many other complications, and you see that there is not going to be a simple law. On the other hand, it is a remarkable fact that the drag force on an airplane is approximately a constant times the square of the velocity, or $F \sim cv^2$.

Now what is the status of such a law, is it analogous to $F = ma$? Not at all, because in the first place this law is an empirical thing that is obtained roughly by tests in a wind tunnel. You say, "Well $F = ma$ might be empirical too." That is not the reason that there is a difference. The difference is not that it is empirical, but that, as we understand nature, this law is the result of an enormous complexity of events and is not, fundamentally, a simple thing. If we continue to study it more and more, measuring more and more accurately, the law will continue to become *more* complicated, not *less*. In other words, as we study this law of the drag on an airplane more and more closely, we find out that it is "falser" and "falser," and the more deeply we study it, and the more accurately we measure, the more complicated the truth becomes; so in that sense we consider it not to result from a simple, fundamental process, which agrees with our original surmise. For example, if the velocity is extremely low, so low that an ordinary airplane is not flying, as when the airplane is dragged slowly through the air, then the law changes, and the drag friction depends more nearly linearly on the velocity. To take another example, the frictional drag on a ball or a bubble or anything that is moving slowly through a viscous liquid like honey, is proportional to the velocity, but for motion so fast that the fluid swirls around (honey does not but water and air do) then the drag becomes more nearly proportional to the square of the velocity ($F = cv^2$), and if the velocity continues to increase, then even this law begins to fail. People who say, "Well the coefficient changes slightly," are dodging the issue. Second, there are other great complications: can this force on the airplane be divided or analyzed as a force on the wings, a force on the front, and so on? Indeed, this can be done, if we are concerned about the torques here and there, but then we have to get special laws for the force on the wings, and so on. It is an amazing fact that the force on a wing depends upon the other wing: in other words, if we take the airplane apart and put just one wing in the air, then the force is not the same as if the rest of the plane were there. The reason, of course, is that some of the wind that hits the front goes around to the wings and changes the force on the wings. It seems a miracle that there is such a simple, rough, empirical law that can be used in the design of airplanes, but this law is not in the same class as the *basic* laws of physics, and further study of it will only make it more and more complicated. A study of how the coefficient c depends on the shape of the front of the airplane is, to put it mildly, frustrating. There just is no simple law for determining the coefficient in terms of the shape of the airplane. In contrast, the law of gravitation is simple, and further study only indicates its greater simplicity.

We have just discussed two cases of friction, resulting from fast movement in air and slow movement in honey. There is another kind of friction, called dry friction or sliding friction, which occurs when one solid body slides on another. In this case a force is needed to maintain motion. This is called a frictional force, and its origin, also, is a very complicated matter. Both surfaces of contact are irregular, on an atomic level. There are many points of contact where the atoms seem to cling together, and then, as the sliding body is pulled along, the atoms snap apart and vibration ensues; something like that has to happen. Formerly the mechanism of this friction was thought to be very simple, that the surfaces were merely full of irregularities and the friction originated in lifting the slider over the bumps; but this cannot be, for there is no loss of energy in that process, whereas power is in fact consumed. The mechanism of power loss is that as the slider snaps over the bumps, the bumps deform and then generate waves and atomic motions and, after a while, heat, in the two bodies. Now it is very remarkable that again, empirically, this friction can be described approximately by a simple law. This law is that the force needed to overcome friction and to drag one object over another depends upon the normal force (i.e., perpendicular to the surface) between the two surfaces that are in contact. Actually, to a fairly good approximation, the frictional force is proportional to this normal force, and has a more or less constant coefficient; that is,

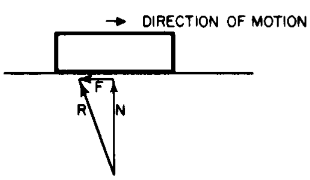

Fig. 12–1. The relation between frictional force and the normal force for sliding contact.

$$F = \mu N, \tag{12.1}$$

where μ is called the *coefficient of friction* (Fig. 12–1). Although this coefficient is not exactly constant, the formula is a good empirical rule for judging approximately the amount of force that will be needed in certain practical or engineering circumstances. If the normal force or the speed of motion gets too big, the law fails because of the excessive heat generated. It is important to realize that each of these empirical laws has its limitations, beyond which it does not really work.

That the formula $F = \mu N$ is approximately correct can be demonstrated by a simple experiment. We set up a plane, inclined at a small angle θ, and place a block of weight W on the plane. We then tilt the plane at a steeper angle, until the block just begins to slide from its own weight. The component of the weight downward along the plane is $W \sin \theta$, and this must equal the frictional force F when the block is sliding uniformly. The component of the weight normal to the plane is $W \cos \theta$, and this is the normal force N. With these values, the formula becomes $W \sin \theta = \mu W \cos \theta$, from which we get $\mu = \sin \theta / \cos \theta = \tan \theta$. If this law were exactly true, an object would start to slide at some definite inclination. If the same block is loaded by putting extra weight on it, then, although W is increased, all the forces in the formula are increased in the same proportion, and W cancels out. If μ stays constant, the loaded block will slide again at the same slope. When the angle θ is determined by trial with the original weight, it is found that with the greater weight the block will slide at about the same angle. This will be true even when one weight is many times as great as the other, and so we conclude that the coefficient of friction is independent of the weight.

In performing this experiment it is noticeable that when the plane is tilted at about the correct angle θ, the block does not slide steadily but in a halting fashion. At one place it may stop, at another it may move with acceleration. This behavior indicates that the coefficient of friction is only roughly a constant, and varies from place to place along the plane. The same erratic behavior is observed whether the block is loaded or not. Such variations are caused by different degrees of smoothness or hardness of the plane, and perhaps dirt, oxides, or other foreign matter. The tables that list purported values of μ for "steel on steel," "copper on copper," and the like, are all false, because they ignore the factors mentioned above, which really determine μ. The friction is never due to "copper on copper," etc., but to the impurities clinging to the copper.

In experiments of the type described above, the friction is nearly independent of the velocity. Many people believe that the friction to be overcome to get something started (static friction) exceeds the force required to keep it sliding (sliding friction), but with dry metals it is very hard to show any difference. The opinion probably arises from experiences where small bits of oil or lubricant are present, or where blocks, for example, are supported by springs or other flexible supports so that they appear to bind.

It is quite difficult to do accurate quantitative experiments in friction, and the laws of friction are still not analyzed very well, in spite of the enormous engineering value of an accurate analysis. Although the law $F = \mu N$ is fairly accurate once the surfaces are standardized, the reason for this form of the law is not really under-understood. To show that the coefficient μ is nearly independent of velocity requires some delicate experimentation, because the apparent friction is much reduced if the lower surface vibrates very fast. When the experiment is done at very high speed, care must be taken that the objects do not vibrate relative to one another, since apparent decreases of the friction at high speed are often due to vibrations. At any rate, this friction law is another of those semiempirical laws that are not thoroughly understood, and in view of all the work that has been done it is surprising that more understanding of this phenomenon has not come about. At the present time, in fact, it is impossible even to estimate the coefficient of friction between two substances.

It was pointed out above that attempts to measure μ by sliding pure substances such as copper on copper will lead to spurious results, because the surfaces in contact are not pure copper, but are mixtures of oxides and other impurities. If we try to get absolutely pure copper, if we clean and polish the surfaces, outgas the materials in a vacuum, and take every conceivable precaution, we still do not get μ. For if we tilt the apparatus even to a vertical position, the slider will not fall off—the two pieces of copper stick together! The coefficient μ, which is ordinarily less than unity for reasonably hard surfaces, becomes several times unity! The reason for this unexpected behavior is that when the atoms in contact are all of the same kind, there is no way for the atoms to "know" that they are in different pieces of copper. When there are other atoms, in the oxides and greases and more complicated thin surface layers of contaminants in between, the atoms "know" when they are not on the same part. When we consider that it is forces between atoms that hold the copper together as a solid, it should become clear that it is impossible to get the right coefficient of friction for pure metals.

The same phenomenon can be observed in a simple home-made experiment with a flat glass plate and a glass tumbler. If the tumbler is placed on the plate and pulled along with a loop of string, it slides fairly well and one can feel the coefficient of friction; it is a little irregular, but it is a coefficient. If we now wet the glass plate and the bottom of the tumbler and pull again, we find that it binds, and if we look closely we shall find scratches, because the water is able to lift the grease and the other contaminants off the surface, and then we really have a glass-to-glass contact; this contact is so good that it holds tight and resists separation so much that the glass is torn apart; that is, it makes scratches.

12–3 Molecular forces

We shall next discuss the characteristics of molecular forces. These are forces between the atoms, and are the ultimate origin of friction. Molecular forces have never been satisfactorily explained on a basis of classical physics; it takes quantum mechanics to understand them fully. Empirically, however, the force between atoms is illustrated schematically in Fig. 12–2, where the force F between two atoms is plotted as a function of the distance r between them. There are different cases: in the water molecule, for example, the negative charges sit more on the oxygen, and the mean positions of the negative charges and of the positive charges are not at the same point; consequently, another molecule nearby feels a relatively large force, which is called a dipole-dipole force. However, for many systems the charges are very much better balanced, in particular for oxygen gas, which is perfectly symmetrical. In this case, although the minus charges and the plus charges are dispersed over the molecule, the distribution is such that the center of the minus charges and the center of the plus charges coincide. A molecule where the centers do not coincide is called a polar molecule, and charge times the separation between centers is called the dipole moment. A nonpolar molecule is one in which the centers of the charges coincide. For all nonpolar molecules, in which all the electrical forces are neutralized, it nevertheless turns out that the force at very large distances is an attraction and varies inversely as the seventh power of the distance, or $F = k/r^7$, where k is a constant that depends on the molecules. Why this is we shall learn only when we learn quantum mechanics. When there are dipoles the forces are greater. When atoms or molecules get too close they repel with a very large repulsion; that is what keeps us from falling through the floor!

These molecular forces can be demonstrated in a fairly direct way: one of these is the friction experiment with a sliding glass tumbler; another is to take two very carefully ground and lapped surfaces which are very accurately flat, so that the surfaces can be brought very close together. An example of such surfaces is the Johansson blocks that are used in machine shops as standards for making accurate length measurements. If one such block is slid over another very carefully and the upper one is lifted, the other one will adhere and also be lifted by the molecular forces, exemplifying the direct attraction between the atoms on one block for the atoms on the other block.

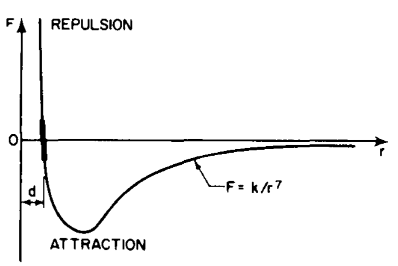

Fig. 12–2. The force between two atoms as a function of their distance of separation.

Nevertheless these molecular forces of attraction are still not fundamental in the sense that gravitation is fundamental; they are due to the vastly complex interactions of all the electrons and nuclei in one molecule with all the electrons and nuclei in another. Any simple-looking formula we get represents a summation of complications, so we still have not got the fundamental phenomena.

Since the molecular forces attract at large distances and repel at short distances, as shown in Fig. 12–2, we can make up solids in which all the atoms are held together by their attractions and held apart by the repulsion that sets in when they are too close together. At a certain distance d (where the graph in Fig. 12–2 crosses the axis) the forces are zero, which means that they are all balanced, so that the molecules stay that distance apart from one another. If the molecules are pushed closer together than the distance d they all show a repulsion, represented by the portion of the graph above the r-axis. To push the molecules only slightly closer together requires a great force, because the molecular repulsion rapidly becomes very great at distances less than d. If the molecules are pulled slightly apart there is a slight attraction, which increases as the separation increases. If they are pulled sufficiently hard, they will separate permanently—the bond is broken.

If the molecules are pushed only a *very small* distance closer, or pulled only a *very small* distance farther than d, the corresponding distance along the curve of Fig. 12–2 is also very small, and can then be approximated by a straight line. Therefore, in many circumstances, if the displacement is not too great *the force is proportional to the displacement*. This principle is known as Hooke's law, or the law of elasticity, which says that the force in a body which tries to restore the body to its original condition when it is distorted is proportional to the distortion. This law, of course, holds true only if the distortion is relatively small; when it gets too large the body will be torn apart or crushed, depending on the kind of distortion. The amount of force for which Hooke's law is valid depends upon the material; for instance, for dough or putty the force is very small, but for steel it is relatively large. Hooke's law can be nicely demonstrated with a long coil spring, made of steel and suspended vertically. A suitable weight hung on the lower end of the spring produces a tiny twist throughout the length of the wire, which results in a small vertical deflection in each turn and adds up to a large displacement if there are many turns. If the total elongation produced, say, by a 100-gram weight, is measured, it is found that additional weights of 100 grams will each produce an additional elongation that is very nearly equal to the stretch that was measured for the first 100 grams. This constant ratio of force to displacement begins to change when the spring is overloaded, i.e., Hooke's law no longer holds.

12–4 Fundamental forces. Fields

We shall now discuss the only remaining forces that are fundamental. We call them fundamental in the sense that their laws are fundamentally simple. We shall first discuss electrical force. Objects carry electrical charges which consist simply of electrons or protons. If any two bodies are electrically charged, there is an electrical force between them, and if the magnitudes of the charges are q_1 and q_2, respectively, the force varies inversely as the square of the distance between the charges, or $F = (\text{const})\, q_1q_2/r^2$. For unlike charges, this law is like the law of gravitation, but for *like* charges the force is repulsive and the sign (direction) is reversed. The charges q_1 and q_2 can be intrinsically either positive or negative, and in any specific application of the formula the direction of the force will come out right if the q's are given the proper plus or minus sign; the force is directed along the line between the two charges. The constant in the formula depends, of course, upon what units are used for the force, the charge, and the distance. In current practice the charge is measured in coulombs, the distance in meters, and the force in newtons. Then, in order to get the force to come out properly in newtons, the constant (which for historical reasons is written $1/4\pi\epsilon_0$) takes the numerical value

$$\epsilon_0 = 8.854 \times 10^{-12}\,\text{coul}^2/\text{newton} \cdot \text{m}^2$$

or

$$1/4\pi\epsilon_0 = 8.99 \times 10^9\,\text{n} \cdot \text{m}^2/\text{coul}^2.$$

Thus the force law for static charges is

$$\mathbf{F} = q_1q_2\mathbf{r}/4\pi\epsilon_0 r^3. \tag{12.2}$$

In nature, the most important charge of all is the charge on a single electron, which is 1.60×10^{-19} coulomb. In working with electrical forces between fundamental particles rather than with large charges, many people prefer the combination $(q_{\text{el}})^2/4\pi\epsilon_0$, in which q_{el} is defined as the charge on an electron. This combination occurs frequently, and to simplify calculations it has been defined by the symbol e^2; its numerical value in the mks system of units turns out to be $(1.52 \times 10^{-14})^2$. The advantage of using the constant in this form is that the force between two electrons in newtons can then be written simply as e^2/r^2, with r in meters, without all the individual constants. Electrical forces are much more complicated than this simple formula indicates, since the formula gives the force between two objects only when the objects are standing still. We shall consider the more general case shortly.

In the analysis of forces of the more fundamental kinds (not such forces as friction, but the electrical force or the gravitational force), an interesting and very important concept has been developed. Since at first sight the forces are very much more complicated than is indicated by the inverse-square laws and these laws hold true only when the interacting bodies are standing still, an improved method is needed to deal with the very complex forces that ensue when the bodies start to move in a complicated way. Experience has shown that an approach known as the concept of a "field" is of great utility for the analysis of forces of this type. To illustrate the idea for, say, electrical force, suppose we have two electrical charges, q_1 and q_2, located at points P and R respectively. Then the force between the charges is given by

$$\mathbf{F} = q_1 q_2 \mathbf{r}/r^3. \tag{12.3}$$

To analyze this force by means of the field concept, we say that the charge q_1 at P produces a "condition" at R, such that when the charge q_2 is placed at R it "feels" the force. This is one way, strange perhaps, of describing it; we say that the force $\mathbf{F}$ on q_2 at R can be written in two parts. It is q_2 multiplied by a quantity $\mathbf{E}$ that would be there whether q_2 were there or not (provided we keep all the other charges in their right places). $\mathbf{E}$ is the "condition" produced by q_1, we say, and $\mathbf{F}$ is the response of q_2 to $\mathbf{E}$. $\mathbf{E}$ is called an *electric field*, and it is a vector. The formula for the electric field $\mathbf{E}$ that is produced at R by a charge q_1 at P is the charge q_1 times the constant $1/4\pi\epsilon_0$ divided by r^2 (r is the distance from P to R), and it is acting in the direction of the radius vector (the radius vector $\mathbf{r}$ divided by its own length). The expression for $\mathbf{E}$ is thus

$$\mathbf{E} = q_1 \mathbf{r}/4\pi\epsilon_0 r^3. \tag{12.4}$$

We then write

$$\mathbf{F} = q_2\,\mathbf{E}, \tag{12.5}$$

which expresses the force, the field, and the charge in the field. What is the point of all this? The point is to divide the analysis into two parts. One part says that something *produces* a field. The other part says that something is *acted on* by the field. By allowing us to look at the two parts independently, this separation of the analysis simplifies the calculation of a problem in many situations. If many charges are present, we first work out the total electric field produced at R by all the charges, and then, knowing the charge that is placed at R, we find the force on it.

In the case of gravitation, we can do exactly the same thing. In this case, where the force $\mathbf{F} = -Gm_1m_2\mathbf{r}/r^3$, we can make an analogous analysis, as follows: the force on a body in a gravitational field is the mass of that body times the field $\mathbf{C}$. The force on m_2 is the mass m_2 times the field $\mathbf{C}$ produced by m_1; that is, $\mathbf{F} = m_2\mathbf{C}$. Then the field $\mathbf{C}$ produced by a body of mass m_1 is $\mathbf{C} = -Gm_1\mathbf{r}/r^3$ and it is directed radially, as in the electrical case.

In spite of how it might at first seem, this separation of one part from another is not a triviality. It would be trivial, just another way of writing the same thing, if the laws of force were simple, but the laws of force are so complicated that it turns out that the fields have a reality that is almost independent of the objects which create them. One can do something like shake a charge and produce an effect, a field, at a distance; if one then stops moving the charge, the field keeps track of all the past, because the interaction between two particles is not instantaneous. It is desirable to have some way to remember what happened previously. If the force upon some charge depends upon where another charge was yesterday, which it does, then we need machinery to keep track of what went on yesterday, and that is the character of a field. So when the forces get more complicated, the field becomes more and more real, and this technique becomes less and less of an artificial separation.

In analyzing forces by the use of fields, we need two kinds of laws pertaining to fields. The first is the response to a field, and that gives the equations of motion. For example, the law of response of a mass to a gravitational field is that the force is equal to the mass times the gravitational field; or, if there is also a charge on the body, the response of the charge to the electric field equals the charge times the electric field. The second part of the analysis of nature in these situations is to formulate the laws which determine the strength of the field and how it is produced. These laws are sometimes called the *field equations*. We shall learn more about them in due time, but shall write down a few things about them now.

First, the most remarkable fact of all, which is true exactly and which can be easily understood, is that the total electric field produced by a number of sources is the vector sum of the electric fields produced by the first source, the second source, and so on. In other words, if we have numerous charges making a field, and if all by itself one of them would make the field $\mathbf{E}_1$, another would make the field $\mathbf{E}_2$, and so on, then we merely add the vectors to get the total field. This principle can be expressed as

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \cdots \tag{12.6}$$

or, in view of the definition given above,

$$\mathbf{E} = \sum_i \frac{q_i \mathbf{r}_i}{4\pi\epsilon_0 r_i^3}. \tag{12.7}$$

Can the same methods be applied to gravitation? The force between two masses m_1 and m_2 was expressed by Newton as $\mathbf{F} = Gm_1m_2\mathbf{r}/r^3$. But according to the field concept, we may say that m_1 creates a field $\mathbf{C}$ in all the surrounding space, such that the force on m_2 is given by

$$\mathbf{F} = m_2\mathbf{C}. \tag{12.8}$$

By complete analogy with the electrical case,

$$\mathbf{C} = -Gm_i\mathbf{r}_i/r_i^3 \tag{12.9}$$

and the gravitational field produced by several masses is

$$\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2 + \mathbf{C}_3 + \cdots \tag{12.10}$$

In Chapter 7, in working out a case of planetary motion, we used this principle in essence. We simply added all the force vectors to get the resultant force on a planet. If we divide out the mass of the planet in question, we get Eq. (12.10).

Equations (12.6) and (12.10) express what is known as *the principle of superposition* of fields. This principle states that the total field due to all the sources is the sum of the fields due to each source. So far as we know today, for electricity this is an absolutely guaranteed law, which is true even when the force law is complicated because of the motions of the charges. There are apparent violations, but more careful analysis has always shown these to be due to the overlooking of certain moving charges. However, although the principle of superposition applies exactly for electrical forces, it is not exact for gravity if the field is too strong, and Newton's equation (12.10) is only approximate, according to Einstein's gravitational theory.

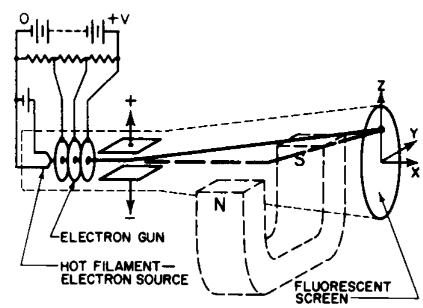

Fig. 12–3. An electron-beam tube.

Closely related to electrical force is another kind, called magnetic force, and this too is analyzed in terms of a field. Some of the qualitative relations between electrical and magnetic forces can be illustrated by an experiment with an electron-ray tube (Fig. 12–3). At one end of such a tube is a source that emits a stream of electrons. Within the tube are arrangements for accelerating the electrons to a high speed and sending some of them in a narrow beam to a fluorescent screen at the other end of the tube. A spot of light glows in the center of the screen where the electrons strike, and this enables us to trace the electron path. On the way to the screen the electron beam passes through a narrow space between a pair of parallel metal plates, which are arranged, say, horizontally. A voltage can be applied across the plates, so that either plate can be made negative at will. When such a voltage is present, there is an electric field between the plates.

The first part of the experiment is to apply a negative voltage to the lower plate, which means that extra electrons have been placed on the lower plate. Since like charges repel, the light spot on the screen instantly shifts upward. (We could also say this in another way—that the electrons "felt" the field, and responded by deflecting upward.) We next reverse the voltage, making the *upper* plate negative. The light spot on the screen now jumps below the center, showing that the electrons in the beam were repelled by those in the plate above them. (Or we could say again that the electrons had "responded" to the field, which is now in the reverse direction.)

The second part of the experiment is to disconnect the voltage from the plates and test the effect of a magnetic field on the electron beam. This is done by means of a horseshoe magnet, whose poles are far enough apart to more or less straddle the tube. Suppose we hold the magnet below the tube in the same orientation as the letter U, with its poles up and part of the tube in between. We note that the light spot is deflected, say, upward, as the magnet approaches the tube from below. So it appears that the magnet repels the electron beam. However, it is not that simple, for if we invert the magnet without reversing the poles side-for-side, and now approach the tube from above, the spot still moves *upward*, so the electron beam is *not* repelled; instead, it appears to be attracted this time. Now we start again, restoring the magnet to its original U orientation and holding it below the tube, as before. Yes, the spot is still deflected upward; but now turn the magnet 180 degrees around a vertical axis, so that it is still in the U position but the poles are reversed side-for-side. Behold, the spot now jumps downward, and stays down, even if we invert the magnet and approach from above, as before.

To understand this peculiar behavior, we have to have a new combination of forces. We explain it thus: Across the magnet from one pole to the other there is a *magnetic field.* This field has a direction which is always away from one particular pole (which we could mark) and toward the other. Inverting the magnet did not change the direction of the field, but reversing the poles side-for-side did reverse its direction. For example, if the electron velocity were horizontal in the x-direction and the magnetic field were also horizontal but in the y-direction, the magnetic force *on the moving electrons* would be in the z-direction, i.e., up or down, depending on whether the field was in the positive or negative y-direction.

Although we shall not at the present time give the correct law of force between charges moving in an arbitrary manner, one relative to the other, because it is too complicated, we shall give one aspect of it: the complete law of the forces *if the fields are known.* The force on a charged object depends upon its motion; if, when the object is standing still at a given place, there is some force, this is taken to be proportional to the charge, the coefficient being what we call the *electric field.* When the object moves the force may be different, and the correction, the new "piece" of force, turns out to be dependent exactly *linearly on the velocity*, but at *right angles* to $\mathbf{v}$ and to another vector quantity which we call the *magnetic induction* $\mathbf{B}$. If the components of the electric field $\mathbf{E}$ and the magnetic induction $\mathbf{B}$ are, respectively, (E_x, E_y, E_z) and (B_x, B_y, B_z), and if the velocity $\mathbf{v}$ has the components (v_x, v_y, v_z), then the total electric and magnetic force on a moving charge q has the components

$$\begin{aligned} F_x &= q(E_x + v_yB_z - v_zB_y),\\ F_y &= q(E_y + v_zB_x - v_xB_z),\\ F_z &= q(E_z + v_xB_y - v_yB_x). \end{aligned} \tag{12.11}$$

If, for instance, the only component of the magnetic field were B_y and the only component of the velocity were v_x, then the only term left in the magnetic force would be a force in the z-direction, at right angles to both $\mathbf{B}$ and $\mathbf{v}$.

12–5 Pseudo forces

The next kind of force we shall discuss might be called a pseudo force. In Chapter 11 we discussed the relationship between two people, Joe and Moe, who use different coordinate systems. Let us suppose that the positions of a particle as measured by Joe are x and by Moe are x'; then the laws are as follows:

$$x = x' + s, \qquad y = y', \qquad z = z',$$

where s is the displacement of Moe's system relative to Joe's. If we suppose that the laws of motion are correct for Joe, how do they look for Moe? We find first, that

$$dx/dt = dx'/dt + ds/dt.$$

Previously, we considered the case where s was constant, and we found that s made no difference in the laws of motion, since $ds/dt = 0$; ultimately, therefore, the laws of physics were the same in both systems. But another case we can take is that $s = ut$, where u is a uniform velocity in a straight line. Then s is not constant, and ds/dt is not zero, but is u, a constant. However, the acceleration d^2x/dt^2 is still the same as d^2x'/dt^2, because $du/dt = 0$. This proves the law that we used in Chapter 10, namely, that if we move in a straight line with uniform velocity the laws of physics will look the same to us as when we are standing still. That is the Galilean transformation. But we wish to discuss the interesting case where s is still more complicated, say $s = at^2/2$. Then $ds/dt = at$ and $d^2s/dt^2 = a$, a uniform acceleration; or in a still more complicated case, the acceleration might be a function of time. This means that although the laws of force from the point of view of Joe would look like

$$m\frac{d^2x}{dt^2} = F_x,$$

the laws of force as looked upon by Moe would appear as

$$m\frac{d^2x'}{dt^2} = F_{x'} - ma.$$

That is, since Moe's coordinate system is accelerating with respect to Joe's, the extra term ma comes in, and Moe will have to correct his forces by that amount in order to get Newton's laws to work. In other words, here is an apparent, mysterious new force of unknown origin which arises, of course, because Moe has the wrong coordinate system. This is an example of a pseudo force; other examples occur in coordinate systems that are *rotating*.

Another example of pseudo force is what is often called "centrifugal force." An observer in a rotating coordinate system, e.g., in a rotating box, will find mysterious forces, not accounted for by any known origin of force, throwing things outward toward the walls. These forces are due merely to the fact that the observer does not have Newton's coordinate system, which is the simplest coordinate system.

Pseudo force can be illustrated by an interesting experiment in which we push a jar of water along a table, with acceleration. Gravity, of course, acts downward on the water, but because of the horizontal acceleration there is also a pseudo force acting horizontally and in a direction opposite to the acceleration. The resultant of gravity and pseudo force makes an angle with the vertical, and during the acceleration the surface of the water will be perpendicular to the resultant force, i.e., inclined at an angle with the table, with the water standing higher in the rearward side of the jar. When the push on the jar stops and the jar decelerates because of friction, the pseudo force is reversed, and the water stands higher in the forward side of the jar (Fig. 12–4).

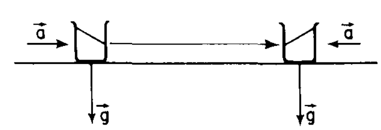

Fig. 12–4. Illustration of a pseudo force

One very important feature of pseudo forces is that they are always proportional to the masses; the same is true of gravity. The possibility exists, therefore, that *gravity itself is a pseudo force*. Is it not possible that perhaps gravitation is due simply to the fact that we do not have the right coordinate system? After all, we can always get a force proportional to the mass if we imagine that a body is accelerating. For instance, a man shut up in a box that is standing still on the earth finds himself held to the floor of the box with a certain force that is proportional to his mass. But if there were no earth at all and the box were standing still, the man inside would float in space. On the other hand, if there were no earth at all and something were *pulling* the box along with an acceleration g, then the man in the box, analyzing physics, would find a pseudo force which would pull him to the floor, just as gravity does.

Einstein put forward the famous hypothesis that accelerations give an imitation of gravitation, that the forces of acceleration (the pseudo forces) *cannot be distinguished* from those of gravity; it is not possible to tell how much of a given force is gravity and how much is pseudo force.

It might seem all right to consider gravity to be a pseudo force, to say that we are all held down because we are accelerating upward, but how about the people in Madagascar, on the other side of the earth—are they accelerating too? Einstein found that gravity could be considered a pseudo force only at one point at a time, and was led by his considerations to suggest that the *geometry of the world* is more complicated than ordinary Euclidean geometry. The present discussion is only qualitative, and does not pretend to convey anything more than the general idea. To give a rough idea of how gravitation could be the result of pseudo forces, we present an illustration which is purely geometrical and does not represent the real situation. Suppose that we all lived in two dimensions, and knew nothing of a third. We think we are on a plane, but suppose we are really on the surface of a sphere. And suppose that we shoot an object along the ground, with no forces on it. Where will it go? It will appear to go in a straight line, but it has to remain on the surface of a sphere, where the shortest distance between two points is along a great circle; so it goes along a great circle. If we shoot another object similarly, but in another direction, it goes along another great circle. Because we think we are on a plane, we expect that these two bodies will continue to diverge linearly with time, but careful observation will show that if they go far enough they move closer together again, as though they were attracting each other. But they are *not* attracting each other—there is just something "weird" about this geometry. This particular illustration does not describe correctly the way in which Euclid's geometry is "weird," but it illustrates that if we distort the geometry sufficiently it is possible that all gravitation is related in some way to pseudo forces; that is the general idea of the Einsteinian theory of gravitation.

12–6 Nuclear forces

We conclude this chapter with a brief discussion of the only other known forces, which are called *nuclear forces.* These forces are within the nuclei of atoms, and although they are much discussed, no one has ever calculated the force between two nuclei, and indeed at present there is no known law for nuclear forces. These forces have a very tiny range which is just about the same as the size of the nucleus, perhaps 10^{-13} centimeter. With particles so small and at such a tiny distance, only the quantum-mechanical laws are valid, not the Newtonian laws. In nuclear analysis we no longer think in terms of forces, and in fact we can replace the force concept with a concept of the energy of interaction of two particles, a subject that will be discussed later. Any formula that can be written for nuclear forces is a rather crude approximation which omits many complications; one might be somewhat as follows: forces within a nucleus do not vary inversely as the square of the distance, but die off exponentially over a certain distance r, as expressed by $F = (1/r^2)\exp(-r/r_0)$, where the distance r_0 is of the order of 10^{-13} centimeter. In other words, the forces disappear as soon as the particles are any great distance apart, although they are very strong within the 10^{-13} centimeter range. So far as they are understood today, the laws of nuclear force are very complex; we do not understand them in any simple way, and the whole problem of analyzing the fundamental machinery behind nuclear forces is unsolved. Attempts at a solution have led to the discovery of numerous strange particles, the π-mesons, for example, but the origin of these forces remains obscure.

13

Work and Potential Energy (A)

13–1 Energy of a falling body

In Chapter 4 we discussed the conservation of energy. In that discussion, we did not use Newton's laws, but it is, of course, of great interest to see how it comes about that energy is in fact conserved in accordance with these laws. For clarity we shall start with the simplest possible example, and then develop harder and harder examples.

The simplest example of the conservation of energy is a vertically falling object, one that moves only in a vertical direction. An object which changes its height under the influence of gravity alone has a kinetic energy T (or K.E.) due to its motion during the fall, and a potential energy mgh, abbreviated U or P.E., whose sum is constant:

$$\underset{\text{KE}}{\tfrac{1}{2}mv^2} + \underset{\text{PE}}{mgh} = \text{const},$$

or

$$T + U = \text{const}. \tag{13.1}$$

Now we would like to show that this statement is true. What do we mean, show it is true? From Newton's Second Law we can easily tell how the object moves, and it is easy to find out how the velocity varies with time, namely, that it increases proportionally with the time, and that the height varies as the square of the time. So if we measure the height from a zero point where the object is stationary, it is no miracle that the height turns out to be equal to the square of the velocity times a number of constants. However, let us look at it a little more closely.

Let us find out *directly* from Newton's Second Law how the kinetic energy should change, by taking the derivative of the kinetic energy with respect to time and then using Newton's laws. When we differentiate $\frac{1}{2}mv^2$ with respect to time, we obtain

$$\frac{dT}{dt} = \frac{d}{dt}\left(\tfrac{1}{2}mv^2\right) = \tfrac{1}{2}m2v\,\frac{dv}{dt} = mv\,\frac{dv}{dt}, \tag{13.2}$$

since m is assumed constant. But from Newton's Second Law, $m(dv/dt) = F$, so that

$$dT/dt = Fv. \tag{13.3}$$

In general, it will come out to be $\mathbf{F}\cdot\mathbf{v}$, but in our one-dimensional case let us leave it as the force times the velocity.

Now in our simple example the force is constant, equal to $-mg$, a vertical force (the minus sign means that it acts downward), and the velocity, of course, is the rate of change of the vertical position, or height h, with time. Thus the rate of change of the kinetic energy is $-mg(dh/dt)$, which quantity, miracle of miracles, is the rate of change of something else! It is the time rate of change of mgh! Therefore, as time goes on, the changes in kinetic energy and in the quantity mgh are equal and opposite, so that the sum of the two quantities remains constant. Q.E.D.

We have shown, from Newton's second law of motion, that energy is conserved for constant forces when we add the potential energy mgh to the kinetic energy $\frac{1}{2}mv^2$. Now let us look into this further and see whether it can be generalized, and thus advance our understanding. Does it work only for a freely falling body, or is it more general? We expect from our discussion of the conservation of energy