17
Spacetime

17.1 The spacetime of Aristotelian physics

From now on, in this book, our attention will be turned from the largely mathematical considerations that have occupied us in earlier chapters, to the actual pictures of the physical world that theory and observation have led us into. Let us begin by trying to understand that arena within which all the phenomena of the physical universe appear to take place: *spacetime*. We shall find that this notion plays a vital role in most of the rest of this book!

We must first ask why 'spacetime'?[1] What is wrong with thinking of space and time separately, rather than attempting to unify these two seemingly very different notions together into one? Despite what appears to be the common perception on this matter, and despite Einstein's quite superb use of this idea in his framing of the general theory of relativity, spacetime was not Einstein's original idea nor, it appears, was he particularly enthusiastic about it when he first heard of it. Moreover, if we look back with hindsight to the magnificent older relativistic insights of Galileo and Newton, we find that they, too, could in principle have gained great benefit from the spacetime perspective.

In order to understand this, let us go much farther back in history and try to see what kind of spacetime structure would have been appropriate for the dynamical framework of Aristotle and his contemporaries. In Aristotelian physics, there is a notion of Euclidean 3-space $\mathbb{E}^3$ to represent physical space, and the points of this space retain their identity from one moment to the next. This is because the state of rest is dynamically preferred, in the Aristotelian scheme, from all other states of motion. We take the attitude that a particular spatial point, at one moment of time, is the *same* spatial point, at a later moment of time, if a particle situated at that point remains at rest from one moment to the next. Our picture of reality is like the screen in a cinema theatre, where a particular point on the screen retains its identity no matter what kinds of vigorous movement might be projected upon it. See Fig. 17.1.

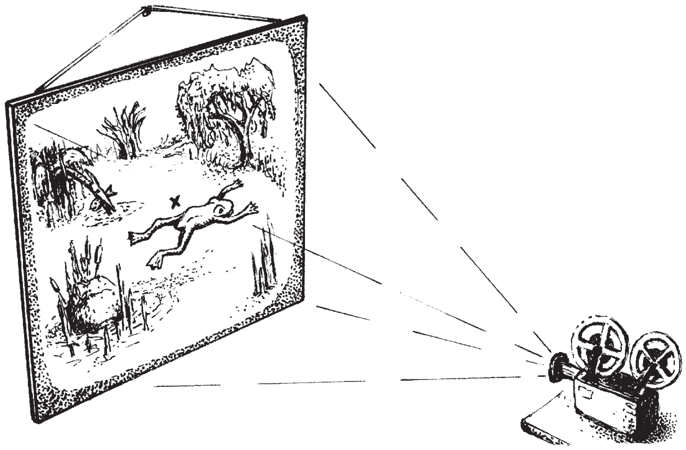

Fig. 17.1 Is physical motion like that perceived on a cinema screen? A particular point on the screen (here marked '×') retains its identity no matter what movement is projected upon it.

Time, also, is represented as a Euclidean space, but as a rather trivial one, namely the 1-dimensional space $\mathbb{E}^1$. Thus, we think of time, as well as physical space, as being a 'Euclidean geometry', rather than as being just a copy of the real line $\mathbb{R}$. This is because $\mathbb{R}$ has a preferred element 0, which would represent the 'zero' of time, whereas in our 'Aristotelian' dynamical view, there is to be no preferred origin. (In this, I am taking an idealized view of what might be called 'Aristotelian dynamics', or 'Aristotelian physics', and I take no viewpoint with regard to what the *actual* Aristotle might have thought!)[2] Had there been a preferred 'origin of time', the dynamical laws could be envisaged as changing when time proceeds away from that preferred origin. With no preferred origin, the laws must remain the same for all time, because there is no preferred *time parameter* which these laws can depend upon.

Likewise, I am taking the view that there is to be no preferred spatial origin, and that space continues indefinitely in all directions, with complete uniformity in the dynamical laws (again, irrespective of what the actual Aristotle might have believed!). In Euclidean geometry, whether 1-dimensional or 3-dimensional, there is a notion of *distance*. In the 3-dimensional spatial case, this is to be ordinary Euclidean distance (measured in metres, or feet, say); in the 1-dimensional case, this distance is the ordinary time interval (measured, say, in seconds).

In Aristotelian physics—and, indeed, in the later dynamical scheme(s) of Galileo and Newton—there is an absolute notion of temporal *simultaneity*. Thus, it has absolute meaning to say, according to such dynamical schemes, that the time here, *at this very moment*, as I sit typing this in my office at home in Oxford, is 'the same time' as some event taking place on the Andromeda galaxy (say the explosion of some supernova star). To return to our analogy of the cinema screen, we can ask whether two projected images, occurring at two widely separated places on the screen, are taking place simultaneously or not. The answer here is clear. The events are to be taken as simultaneous if and only if they occur in the same projected frame. Thus, not only do we have a clear notion of whether or not two (temporally separated) events occur at the same spatial location on the screen, but we also have a clear notion of whether or not two (spatially separated) events occur at the same time. Moreover, if the spatial locations of the two events are different, we have a clear notion of the *distance* between them, whether or not they occur at the same time (i.e. the distance measured along the screen); also, if the times of the two events are different, we have a clear notion of the *time interval* between them, whether or not they occur at the same place.

What this tells us is that, in our Aristotelian scheme, it is appropriate to think of *spacetime* as simply the product

$$\mathcal{A} = \mathbb{E}^1 \times \mathbb{E}^3,$$

which I shall call *Aristotelian spacetime*. This is simply the space of pairs (t, $\mathbf{x}$), where t is an element of $\mathbb{E}^1$, a 'time', and $\mathbf{x}$ is an element of $\mathbb{E}^3$, a 'point in space'. (See Fig. 17.2.) For two different points of $\mathbb{E}^1 \times \mathbb{E}^3$, say ($t$, $\mathbf{x}$) and (t', $\mathbf{x}'$)—i.e. two different *events*—we have a well-defined notion of their spatial separation, namely the distance between the points $\mathbf{x}$ and $\mathbf{x}'$ of $\mathbb{E}^3$, and we also have a well-defined notion of their time difference, namely the separation between t and t' as measured in $\mathbb{E}^1$. In particular, we know whether or not two events occur at the same place (vanishing of spatial displacement) and whether or not they take place at the same time (vanishing of time difference).

17.2 Spacetime for Galilean relativity

Now let us see what notion of spacetime is appropriate for the dynamical scheme introduced by Galileo in 1638. We wish to incorporate the *principle of Galilean relativity* into our spacetime picture. Let us try to

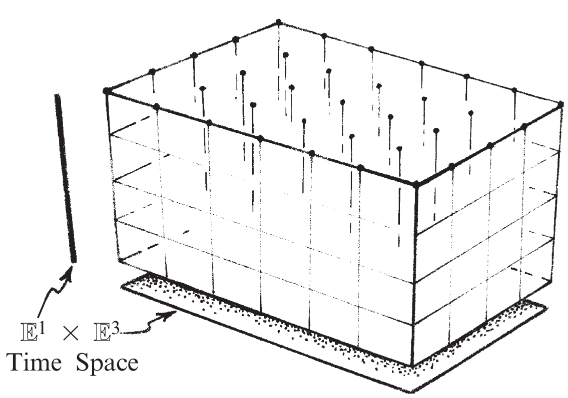

Fig. 17.2 Aristotelian spacetime $\mathcal{A} = \mathbb{E}^1 \times \mathbb{E}^3$ is the space of pairs (t, $\mathbf{x}$), where t ('time') ranges over a Euclidean 1-space $\mathbb{E}^1$, and $\mathbf{x}$ ('point in space') ranges over a Euclidean 3-space $\mathbb{E}^3$.

recall what this principle asserts. It is hard to do better than quote Galileo himself (in a translation due to Stillman Drake[3] which I give here in abbreviated form only; and I strongly recommend an examination of the quote as a whole, for those who have access to it):

> Shut yourself up with some friend in the main cabin below decks on some large ship, and have with you some flies, butterflies, and other small flying animals ... hang up a bottle that empties drop by drop into a wide vessel beneath it ... have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. ... The droplets will fall ... into the vessel beneath without dropping toward the stern, although while the drops are in the air the ship runs many spans ... the butterflies and flies will continue their flights indifferently toward every side, nor will it ever happen that they are concentrated toward the stern, as if tired out from keeping up with the course of the ship. ...

What Galileo teaches us is that the dynamical laws are precisely the same when referred to any uniformly moving frame. (This was an essential ingredient of his wholehearted acceptance of the Copernican scheme, whereby the Earth is allowed to be in motion without our directly noticing this motion, as opposed to its necessarily stationary status according to the earlier Aristotelian framework.) There is nothing to distinguish the physics of the state of rest from that of uniform motion. In terms of what has been said above, what this tells us is that there is no dynamical meaning to saying that a particular point in space is, or is not, the same point as some chosen point in space at a later time. In other words, our cinema-screen analogy is inappropriate! There is no background space—a 'screen'—which remains fixed as time evolves. We cannot meaningfully say that a particular point p in space (say, the point of the exclamation mark on the keyboard of my laptop) is, or is not, the *same* point in space as it was a minute ago. To address this issue more forcefully, consider the rotation of the Earth. According to this motion, a point fixed to the Earth's surface (at the latitude of Oxford, say) will have moved by some 10 miles in the minute under consideration. Accordingly, the point p that I had just selected will now be situated somewhere in the vicinity of the neighbouring town of Witney, or beyond. But wait! I have not taken the Earth's motion about the sun into consideration. If I do that, then I find that p will now be about one hundred times farther off, but in the opposite direction (because it is a little after mid-day, and the Earth's surface, here, now moves oppositely to its motion about the Sun), and the Earth will have moved away from p to such an extent that p is now beyond the reach of the Earth's atmosphere! But should I not have taken into account the sun's motion about the centre of our Milky Way galaxy? Or what about the 'proper motion' of the galaxy itself within the local group? Or the motion of the local group about the centre of the Virgo cluster of which it is a tiny part, or of the Virgo cluster in relation to the vast Coma supercluster, or perhaps the Coma cluster towards 'the Great Attractor'?

Clearly we should take Galileo seriously. There is no meaning to be attached to the notion that any particular point in space a minute from now is to be judged as the *same* point in space as the one that I have chosen. In Galilean dynamics, we do not have just one Euclidean 3-space $\mathbb{E}^3$, as an arena for the actions of the physical world evolving with time, we have a *different* $\mathbb{E}^3$ for each moment in time, with no natural identification between these various $\mathbb{E}^3$s.

It may seem alarming that our very notion of physical space seems to be of something that evaporates completely as one moment passes, and reappears as a completely different space as the next moment arrives! But here the mathematics of Chapter 15 comes to our rescue, for this situation is just the kind of thing that we studied there. *Galilean spacetime* $\mathcal{G}$ is not a product space $\mathbb{E}^1 \times \mathbb{E}^3$, it is a *fibre bundle*[4] with base space $\mathbb{E}^1$ and fibre $\mathbb{E}^3$! In a fibre bundle, there is no pointwise identification between one fibre and the next; nevertheless the fibres fit together to form a connected whole. Each spacetime event is naturally assigned a *time*, as a particular element of one specific 'clock space' $\mathbb{E}^1$, but there is no natural assignment of a spatial location in one specific 'location space' $\mathbb{E}^3$. In the bundle language of §15.2, this natural assignment of a time is achieved by the *canonical projection* from $\mathcal{G}$ to $\mathbb{E}^1$. (See Fig. 17.3; compare also Fig. 15.2.)

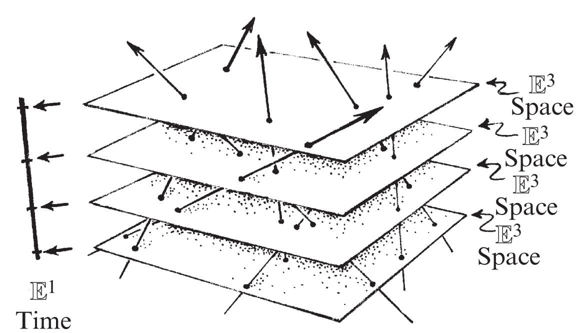

Fig. 17.3 Galilean spacetime $\mathcal{G}$ is fibre bundle with base space $\mathbb{E}^1$ and fibre $\mathbb{E}^3$, so there is no given pointwise identification between different $\mathbb{E}^3$ fibres (no absolute space), whereas each spacetime event is assigned a time via the canonical projection (absolute time). (Compare Fig. 15.2, but the canonical projection to the base is here depicted horizontally.) Particle histories (world lines) are cross-sections of the bundle (compare Fig. 15.6a), the inertial particle motions being depicted here as what $\mathcal{G}$'s structure specifies, that is: 'straight' world lines.

17.3 Newtonian dynamics in spacetime terms

This 'bundle' picture of spacetime is all very well, but how are we to express the *dynamics* of Galileo–Newton in terms of it? It is not surprising that Newton, when he came to formulate his laws of dynamics, found himself driven to a description in which he appeared to favour a notion of 'absolute space'. In fact, Newton was, at least initially, as much of a Galilean relativist as was Galileo himself. This is made clear from the fact that in his original formulation of his laws of motion, he explicitly stated the Galilean principle of relativity as a fundamental law (this being the principle that physical action should be blind to a change from one uniformly moving reference frame to another, the notion of *time* being absolute, as is manifested in the picture above of Galilean spacetime $\mathcal{G}$).

He had originally proposed five (or six) laws, law 4 of which was indeed the Galilean principle,[5] but later he simplified them, in his published *Principia*, to the three 'Newton's laws' that we are now familiar with. For he had realized that these were sufficient for deriving all the others. In order to make the framework for his laws precise, he needed to adopt an 'absolute space' with respect to which his motions were to be described. Had the notion of a 'fibre bundle' been available at the time (admittedly a far-fetched possibility), then it would have been conceivable for Newton to formulate his laws in a way that is completely 'Galilean-invariant'. But without such a notion, it is hard to see how Newton could have proceeded without introducing some concept of 'absolute space', which indeed he did.

What kind of structure must we assign to our 'Galilean spacetime' $\mathcal{G}$? It would certainly be far too strong to endow our fibre bundle $\mathcal{G}$ with a bundle connection (§15.7).[17.1] What we must do, instead, is to provide it with something that is in accordance with *Newton's first law*. This law states that the motion of a particle, upon which no forces act, must be uniform and in a straight line. This is called an *inertial motion*. In spacetime terms, the motion (i.e. 'history') of any particle, whether in inertial motion or not, is represented by a curve, called the *world line* of the particle. In fact, in our Galilean spacetime, world lines must always be *cross-sections* of the Galilean bundle; see §15.3.[17.2] and Fig. 17.3.) The notion of 'uniform and in a straight line', in ordinary spatial terms (an inertial motion), is interpreted simply as 'straight', in spacetime terms. Thus, the Galilean bundle $\mathcal{G}$ must have a structure that encodes the notion of 'straightness' of world lines. One way of saying this is to assert that $\mathcal{G}$ is an *affine* space (§14.1) in which the affine structure, when restricted to individual $\mathbb{E}^3$ fibres, agrees with the Euclidean affine structure of each $\mathbb{E}^3$.

[17.1] Why?

[17.2] Explain the reason for this.

Another way is simply to specify the ∞^6 family of straight lines that naturally resides in $\mathbb{E}^1 \times \mathbb{E}^3$ (the 'Aristotelian' uniform motions) and to take these over to provide the 'straight-line' structure of the Galilean bundle, while 'forgetting' the actual product structure of the Aristotelian spacetime $\mathcal{A}$. (Recall that ∞^6 means a 6-dimensional family; see §16.7.) Yet another way is to assert that the Galilean spacetime, considered as a manifold, possesses a connection which has both vanishing curvature and vanishing torsion (which is quite different from it possessing a bundle connection, when considered as a bundle over $\mathbb{E}^1$).[17.3]

In fact, this third point of view is the most satisfactory, as it allows for the generalizations that we shall be needing in §§17.5,9 in order to describe gravitation in accordance with Einstein's ideas. Having a connection defined on $\mathcal{G}$, we are provided with a notion of *geodesic* (§14.5), and these geodesics (apart from those which are simply straight lines in individual $\mathbb{E}^3$s) define Newton's *inertial motions*. We can also consider world lines that are not geodesics. In ordinary spatial terms, these represent particle motions that accelerate. The actual magnitude of this acceleration is measured, in spacetime terms, as a curvature of the world-line.[17.4] According to *Newton's second law*, this acceleration is equal to the total force on the particle, divided by its mass. (This is Newton's $f = ma$, in the form $a = f \div m$, where a is the particle's acceleration, m is its mass, and f is the total force acting upon it.) Thus, the curvature of a world line, for a particle of given mass, provides a direct measure of the total force acting on that particle.

In standard Newtonian mechanics, the total force on a particle is the (vector) *sum* of contributions from all the other particles (Fig. 17.4a). In any particular $\mathbb{E}^3$ (that is, at any one time), the contribution to the force on one particle, from some other particle, acts in the line joining the two that lies in that particular $\mathbb{E}^3$. That is to say, it acts *simultaneously* between the two particles. (See Fig. 17.4b.) *Newton's third law* asserts that the force on one of these particles, as exerted by the other, is always equal in magnitude and opposite in direction to the force on the other as exerted by the one. In addition, for each different variety of force, there is a *force law*, informing us what function of the spatial distance between the particles the magnitude of that force should be, and what parameters should be used for each type of particle, describing the overall scale for that force. In the particular case of gravity, this function is taken to be the inverse square of the distance, and the overall scale is a certain constant, called Newton's gravitational constant G, multiplied by the product of the two masses

[17.3] Explain these three ways more thoroughly, showing why they all give the same structure.

[17.4] Try to write down an expression for this curvature, in terms of the connection ∇. What normalization condition on the tangent vectors is needed (if any)?

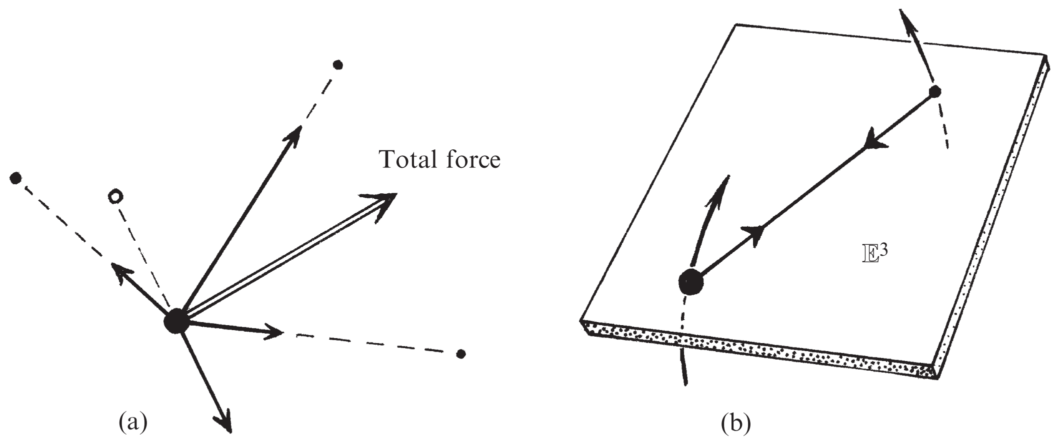

Fig. 17.4 (a) Newtonian force: at any one time, the total force on a particle (double shafted arrow) is the vector sum of contributions (attractive or repulsive) from all other particles. (b) Two particle world lines and the force between them, acting 'instantaneously', in a line joining the two particles, at any one moment, within the particular $\mathbb{E}^3$ that the moment defines. Newton's Third Law asserts that force on one, as exerted by the other, is equal in magnitude and opposite in direction to the force on the other as exerted by the one.

involved. In terms of symbols, we get Newton's well-known formula for the attractive force on a particle of mass m, as exerted by another particle of mass M, a distance r away from it, namely

$$\frac{GmM}{r^2}.$$

It is remarkable that, from just these simple ingredients, a theory of extraordinary power and versatility arises, which can be used with great accuracy to describe the behaviour of macroscopic bodies (and, for most basic considerations, submicroscopic particles also), so long as their speeds are significantly less than that of light. In the case of gravity, the accordance between theory and observation is especially clear, because of the very detailed observations of the planetary motions in our solar system. Newton's theory is now found to be accurate to something like one part in 10^7, which is an extremely impressive achievement, particularly since the accuracy of data that Newton had to go on was only about one ten-thousandth of this (a part in 10^3).

17.4 The principle of equivalence

Despite this extraordinary precision, and despite the fact that Newton's great theory remained virtually unchallenged for nearly two and one half centuries, we now know that this theory is not absolutely precise; more-