special relativity, might not some other method? Much of the literature on the subject of conventionality of simultaneity has been devoted to investigating such alternative methods of synchronizing spatially separated clocks and seeking to reveal definitions equivalent to the setting of a value for ϵ in them. See for example Salmon (1977) to get a clear sense that no such convention-free, alternative method is likely to be found. Note that this literature urges the conventionality of the "one-way" velocity light, that is, the velocity between two spatially separated points. The round trip velocity is not taken to be conventional since only one clock at the common source and destination is needed for its measurement.

We return to the conventionality of simultaneity in Section 5.11 to see one of the most dramatic reversals in debates in the philosophy of space and time. David Malament has recently derived a theorem in special relativity which, he urges, shows that the causal relations of special relativity do *not* leave the simultaneity relation underdetermined and thus the relation cannot be set conventionally within the causal theory of time. He shows that the only nontrivial simultaneity relation definable in terms of the causal relations of special relativity is the familiar standard simultaneity relation of $\epsilon = 1/2$.

Part II: Theories and Methods

The purpose of this part is to introduce the methods now used almost exclusively in recent work in philosophy of space and time. These methods differ from those used in Part I in several important ways.

1. There is less emphasis on theories of a space and time as a set of law-like sentences. Rather the theories are approached semantically (see Chapter 3). Thus the activity of the theorist becomes akin to that of the hobbyist model builder, who seeks to represent a real sailboat by constructing a model that captures as many of its properties as possible. The space and time theorist builds models which are intended to reflect the spatial and temporal properties of reality. However the theorist's models are not constructed out of balsa, glue and string, but out of abstract mathematical entities such as numbers.
2. Theories of space and time—including Newton's theory of space and time—are worked into a spacetime formulation. Thus when Newton's theory is compared with its relativistic rivals, all the theories are formulated in the same manner, ensuring that the differences observed are true differences and not accidents of differing formulations.
3. A major theme of Part I was the separation of the conventional or arbitrary elements of a theory from the factual or, as we now say, "physically significant" elements. A means of effecting automatically this separation is built into the notions of "covariance" and "invariance" to be explained here in Part II.

5.4 A SIMPLE THEORY OF LINEAR TIME

Let us begin by developing a very simple theory of time whose main purpose is to illustrate the use of models and the notions of covariance and invariance in a setting far simpler than the spacetime theories to which we will soon turn. The basic temporal facts of some physically possible world are that it has infinitely many instants, extending indefinitely into the past and future. The set of instants is homogeneous: Every instant is exactly like every other. The set is also assumed to be isotropic: The future and past directions are exactly alike. To capture and make precise these loosely stated facts, let us develop the following sequence of time theories.

5.4.1 The One Coordinate System Formulation

Let us select as the model for our theory the manifold of all real numbers **R**. Each real number in **R** represents a particular instant (see Figure 5.4). This representation relation is a *coordination* of the instants of the physically possible world with the mathematical structure **R** so that the relation is commonly called a *coordinate system*. We can infer many of the temporal properties of the physically possible world from the coordination. For example, the fact that there is no greatest real number represents the fact that there is no last instant, so that the world persists through indefinitely many instants into the future. Similarly, the denseness of **R**—the fact there is always another real number between any two given real numbers—represents the denseness of time. It models the fact that every temporal interval can be divided so that indivisible time "atoms" are disallowed.

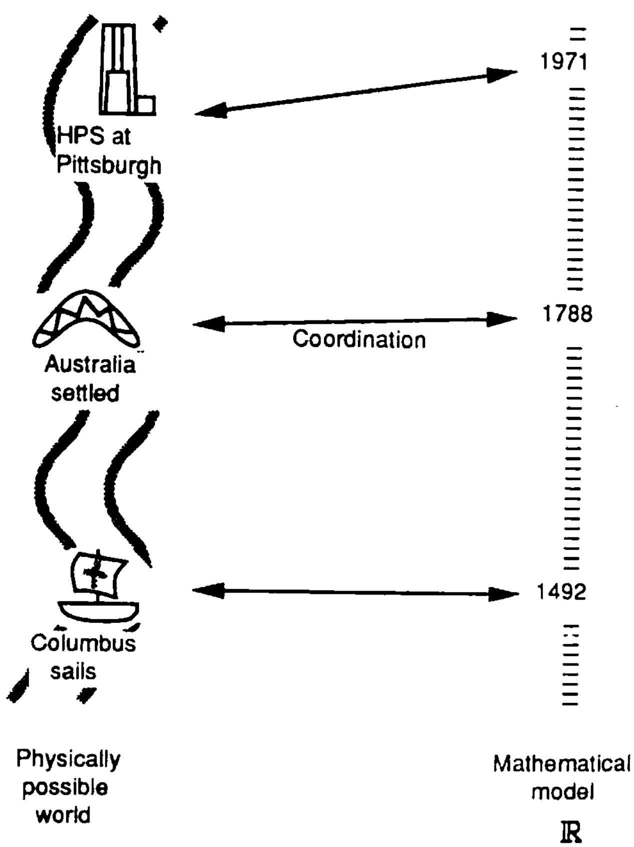


Figure 5.4 Coordinate system for the linear theory of time.

5.4.2 The Standard Formulation

Unfortunately we cannot construe every property of **R** as representing a property of the physically possible world. For example **R** is anisotropic; the direction of increasing real numbers is distinct from that of decreasing real numbers. However, we posited that the physical instants form an isotropic continuum. Similarly, **R** is inhomogeneous; the real number 0 is distinct, for example, from every other number. However, we posited that the physical instants form a homogeneous continuum.

A simple device enables us to designate systematically which are the physically significant properties of the models. To deny physical significance to the anisotropy of **R**, we expand the coordinations of the physically possible world with **R** allowed by the theory. We now allow a new coordination reflected about 0 (see Figure 5.5). Those instants coordinated with 0, 500, 1000, 1500 and so forth in the original coordinate system are now coordinated with 0, −500, −1000, −1500 and so forth in the new system. We call the transformation connecting the two coordinate systems a reflection about 0. If we allow that both the original and reflected coordinate systems are equally good representations of the continuum of physical instants, then the anistropy of **R** no longer enables us to pick out a preferred direction in the continuum of physical instants. The direction picked out by increasing real numbers in one coordinate system is the opposite direction to the one picked out by increasing real numbers in the reflected coordinate system.

Similarly we deny physical significance to the inhomogeneity of **R** by allowing all the coordinate systems produced from the original by a translation of the original coordinate system. For example, in the original coordinate system the instant to which 0 is assigned is singled out as special when compared to the one to which 500 is assigned. We can remove this special status by allowing a second coordinate system in which the latter event is now assigned the value 0. This new coordinate system is produced by translating the original by 500. Figuratively this amounts to "sliding" down by 500 each of the real values coordinated to each instant by the original coordinate system to form the new coordinate system. See Figure 5.5. We ensure that the inhomogeneity of **R** accords no special status to any physical instant by allowing into the theory all coordinate systems produced by a translation from the original by *any* real value. Thus, given any physical instant at all, we can always find a coordinate system in which that instant is assigned the value 0 or, for that matter, any other real value you care to name.

5.4.3 Covariance and Invariance

In sum, the standard formulation of the theory has the original coordinate system as well as all those produced by the coordinate transformations of reflection and translation. Let us call these the *standard* coordinate systems of the theory. The set of reflections and translations form a group of transformations (see the following box) which essentially only means that we never leave the set of transformations if we invert or combine them. It is called the *covariance group* of the theory. Alternatively, we say that the theory is covariant under reflections and translations.

The advantage of the standard formulation over the one coordinate system

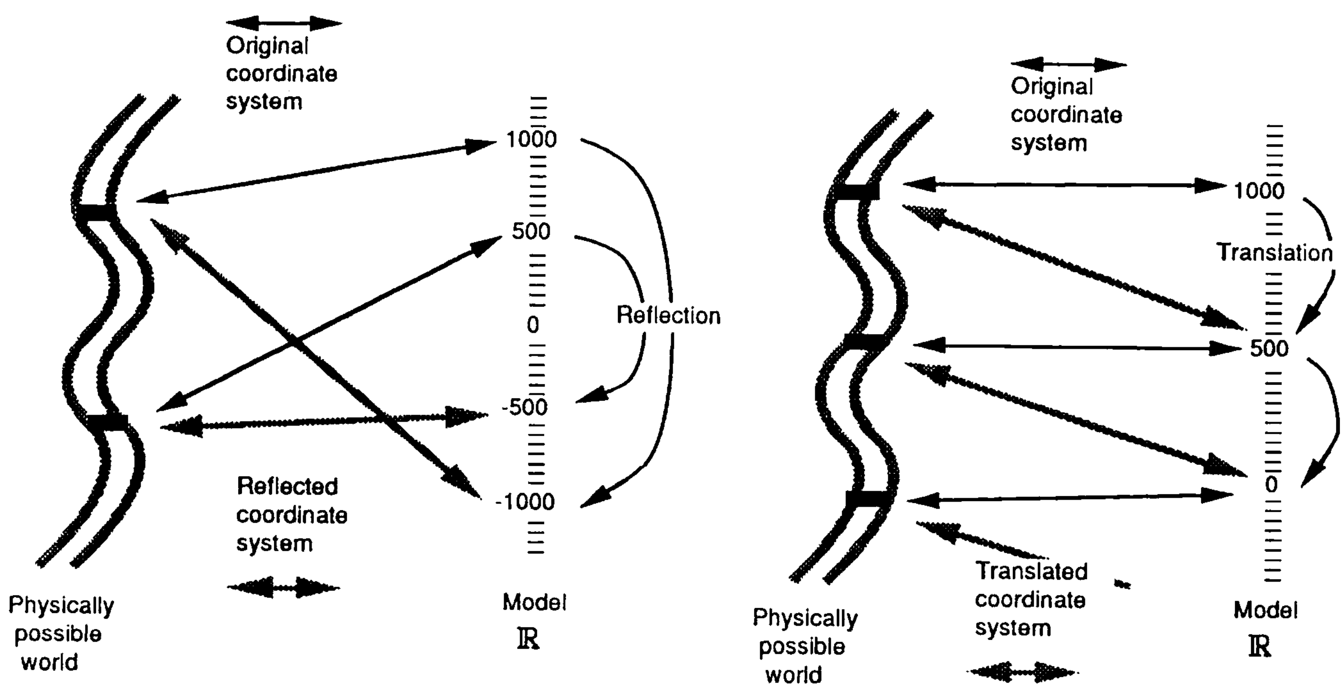


Figure 5.5 Coordinate transformations for the standard formulation of the linear theory to time.

formulation is that it enables us to distinguish the physically significant or factual properties of the theory's model from the arbitrary ones. Those properties are exactly the ones that remain the same in all the coordinate systems of the theory. We can state this important conclusion in another way. By an *invariant* of a transformation, we mean something that remains unchanged under the transformation. Thus we arrive at a principle of paramount importance to all theories of space and time:

> The factual or physically significant quantities of a theory of space and time are the invariants of its covariance group.

All other quantities can be chosen arbitrarily or conventionally. For example, the fact that one coordinate system assigns the real value 27 to some instant is not invariant and thus not physically significant. A different coordinate system will in general assign a different value to the instant. Thus the choice of coordinate system is an arbitrary or conventional stipulation. However if the difference of coordinate values of two instants is 100, then it will be ± 100 in all standard coordinate systems. Thus we conclude that the absolute value of coordinate differences (i.e., the difference as a positive number) in standard coordinate systems is invariant and therefore physically significant. These coordinate differences are interpreted as duration or physical time elapsed, such as might be read by a physical clock.

The strategy of characterizing geometric structure as the invariants of groups has a venerable history. It dates back to Felix Klein's "Erlangen program" of the 1870s in which Klein set out to use the strategy to unify the treatment of the diverse geometries discovered in the nineteenth century.

If the original coordinate system assigns the real value T to some physical instant i, then a new coordinate system produced by a reflection about 0 assigns the new value T' to i where

$$T' = -T$$

and a translation by K assigns the new value T'' to i where

$$T'' = T - K.$$

Combining we can now represent the covariance group of the standard formulation as the set of all transformations given by

$$T^* = At - K$$

where A is $+1$ or -1 and K has any real value. Formally this set of transformations is a group since it satisfies the three conditions:

1. The set contains the identity transformation.
2. Every transformation's inverse is in the set.
3. The composition of two transformations of the set is in the set.

5.4.4 The Generally Covariant Formulation

The adoption of a generally covariant formulation of the theory provides a way of making more explicit just what are the physically significant quantities of the theory. To arrive at the formulation, we expand the allowed coordinate systems to include all those which can be transformed to the original system by smooth invertible transformations on $\mathbb{R}$. Figuratively this means that the allowed transformations of the theory include not just reflections and "slidings" (translations) of the coordinate system but just about any arbitrary "stretching and squeezing" which preserves the smoothness of the coordinate system and the uniqueness of the identification of all instants. However, we cannot leave the theory in this state for we can no longer represent duration by coordinate differences. Coordinate differences are certainly no longer invariant under the arbitrary transformations now allowed. To recover the ability to represent duration invariantly, we must explicitly introduce a new mathematical structure into the theory.

Consider some very small duration between two instants which have coordinate values 1000 and 1001 in a standard coordinate system (see Figure 5.6). The coordinate difference—call it "ΔT"—equals 1 and it is the duration between the two instants. Now introduce a new coordinate system which has been stretched linearly to double the size of the original system, so that to instants originally assigned values 0, 500, 1000, 1001 and so forth are now assigned values 0, 1000, 2000, 2002 and so forth. The coordinate difference in the new system between the same two instants—call it "Δt"—is now equal to 2. To recover the original duration we must multiply the new coordinate difference by a scale factor of 1/2. This scale factor is the extra geometrical structure which we need. Every coordinate system of the generally covariant formulation must be supplied with this scale factor to enable assertions about duration to be made. In general for a *small* duration between two instants whose coordinate values differ by Δt in some coordinate system we have the invariant result:

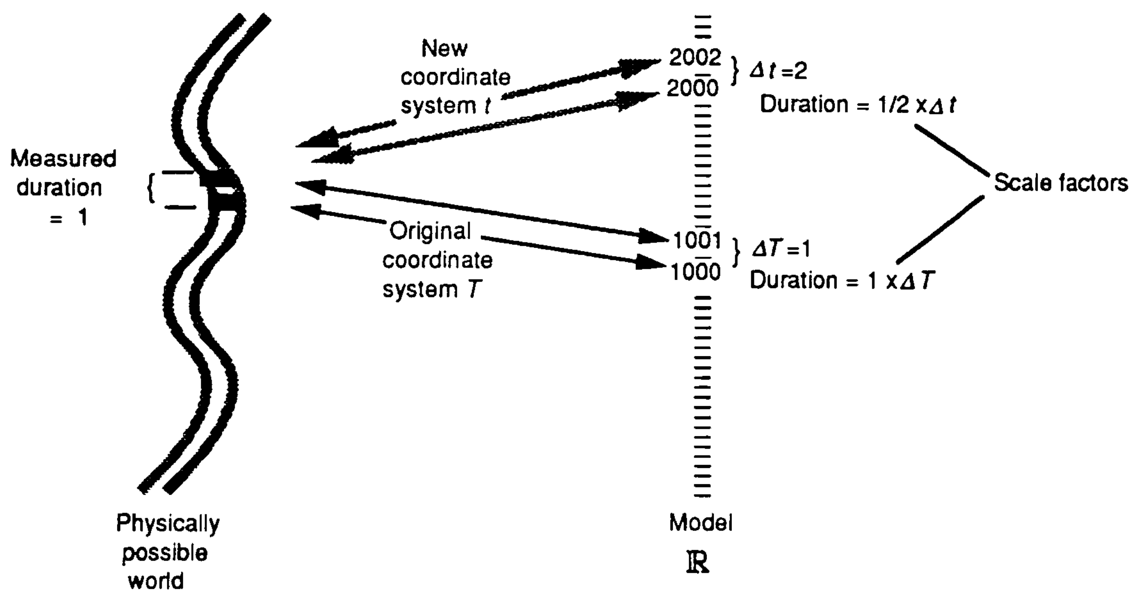


Figure 5.6 Temporal metric for the linear theory of time.

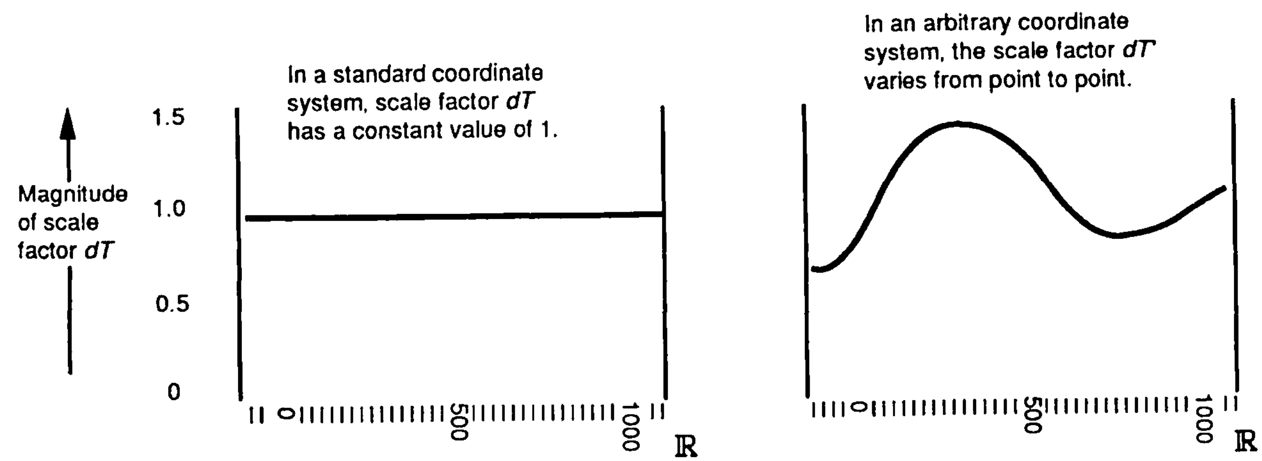


Figure 5.7 Scale factor dT in different coordinate systems.

$$\text{Invariant duration} = \text{Scale factor } X\, |\Delta t|$$

The scale factor of a standard coordinate system is unity. It is 1/2 for the linearly stretched system described here. In some arbitrary coordinate system, the scale factor will have a value that varies from instant to instant according to how much the coordinate system has been stretched or squeezed in the transformation from a standard coordinate system (see Figure 5.7). There is a simple rule—see equation (3) in the following box—for computing how the scale factor will change under an arbitrary coordinate transformation. The existence of such a rule means that the scale factor is a *covariant quantity*: Once we know its value in one coordinate system we invoke its characteristic transformation law to find its value in any other coordinate system. Alternately, such quantities are known as *geometric objects*. See Figure 5.8 for a pictorial representation of this transformation law.

In order to comply with the standard notation, let us represent the scale factor by "dT." The scale factor dT (together with all its transforms) is known as a "covector" or "one-form" and, with regard to its function in the theory, might also be called a "temporal metric" since it is responsible for assigning measurable time

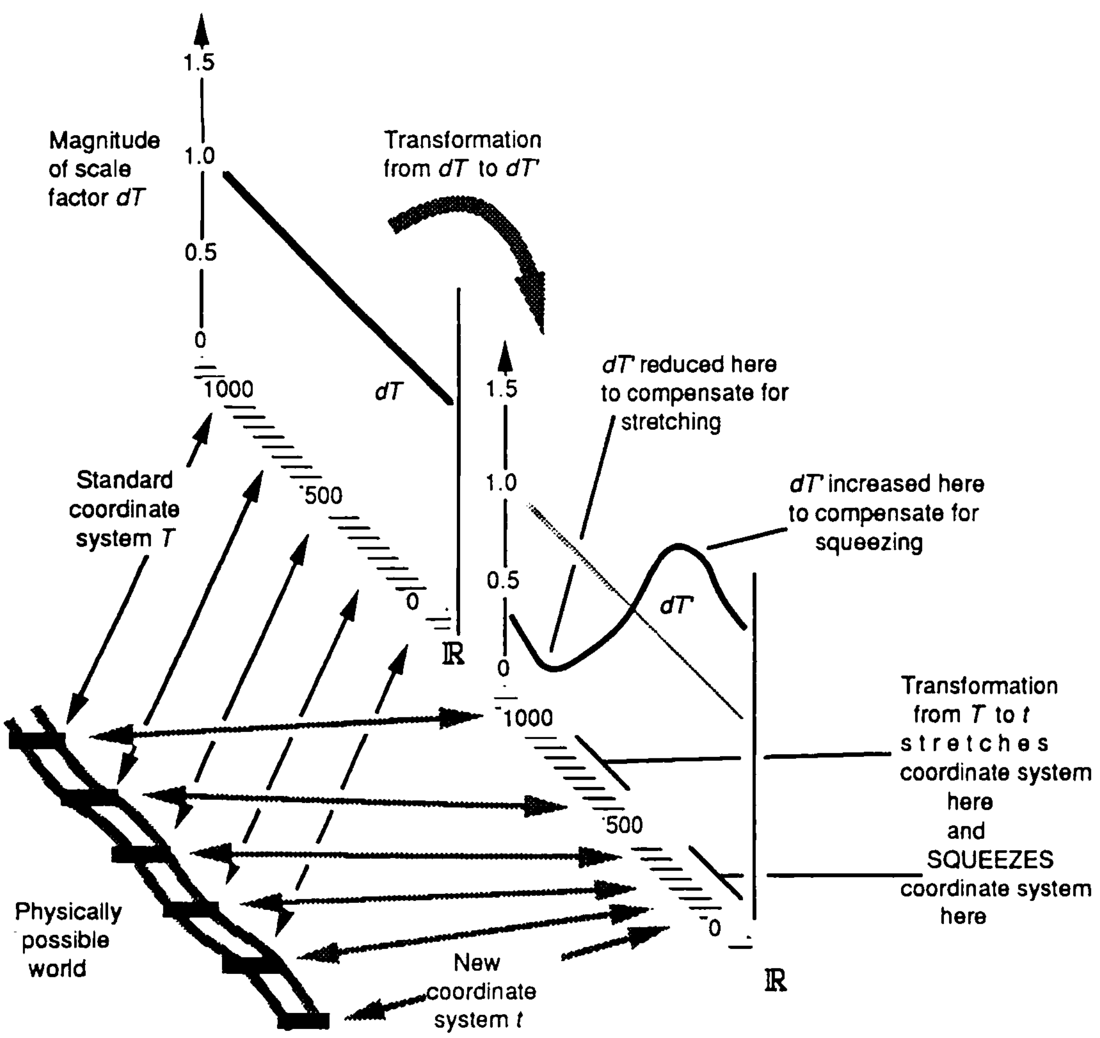


Figure 5.8 Transformation of scale factor dT.

durations to the intervals between instants. (For completeness we note a technical complication. A covector dT assigns positive or negative measures to intervals between instants according to their directions. Because we want no anisotropy in the continuum of instants, only the absolute values of the measures assigned have physical significance.)

The generally covariant formulation of the theory has models of the form $<\mathbf{R},dT>$, where the angle brackets "$<,>$" denote an ordered pair. Every time we change coordinate systems we generate a new scale factor dT. Thus the model set of the theory contains infinitely many models

$$<\mathbf{R},dT>,<\mathbf{R},dT'>,<\mathbf{R},dT''>, \ldots$$

where dT, dT', dT'', . . . can all be transformed into one another and thus represent the one covariant quantity or geometric object. There is a natural division of labor between the two structures of the pair that form the models. The fact that the physical world can be coordinated with **R** gives us its topological properties: Briefly, its instants form a linear continuum with no end points in either direction. Unlike the standard formulation, the coordination with **R** gives us no information on the physically measurable duration between instants. Such information is provided by the temporal metric dT, the second member of the pair.

The model $<\mathbf{R},dT>$ is typical of those used in theories of time, space and spacetime. The models of the theories we now turn to all have the general form

<manifold, geometric object, geometric object, . . . >

The first member of the model, the manifold, represents the topology of the time, space or spacetime in question. Thus it tells us how many dimensions a space has and

---

Let T be a standard coordinate system and ΔT the coordinate difference between two very close instants so that ΔT is also the duration between the instants. We now transform to a new coordinate system t, which need not be a standard coordinate system. We have immediately

$$\text{Duration of interval} = \Delta T = \frac{dT}{dt}\,\Delta t$$

and we can identify dT/dt as the scale factor dT in the coordinate system t. For example, if $t = T^3$, then the scale factor is given by $dT/dt = 1/(3T^2) = 1/(3t^{2/3})$. If we now consider another coordinate system t' with dT' equal to dT/dt', then the chain rule for differentiation, $\frac{dT}{dt'} = \frac{dt}{dt'}\frac{dT}{dt}$, gives us the general transformation law for dT:

$$dT' = \frac{dt}{dt'}\; dT \qquad (3)$$

(3) is the characteristic transformation law for covectors or one-forms.

---

gives us information on its global topology. In the simple linear time theory, time was globally like a line, extending indefinitely into past and future. However, we might want to model a time that is cyclical so that the past and future join. We would then not use **R** as the manifold, but another one-dimensional manifold that is closed like a circle. There are many manifolds more complicated than **R** that the theorist can choose in building models. The remaining members of the model are the geometric objects such as dT that are "painted" onto the canvas of the manifold. They provide the nontopological properties of the space. Thus if we want to know the time elapsed between instants, we look to a temporal metric. In a theory of space, we look to a spatial metric to tell us the distance between two points along some curve. Such a theory is the subject of the next section.

5.5 EUCLIDEAN SPACE

The theory of a Euclidean space is very similar in structure to the linear time theory. Let us consider the case of a two-dimensional Euclidean space. The generalization to the three-dimensional case is entirely straightforward.

The theory's models are built with two-dimensional manifold $\mathbf{R}^2$, where $\mathbf{R}^2$ is

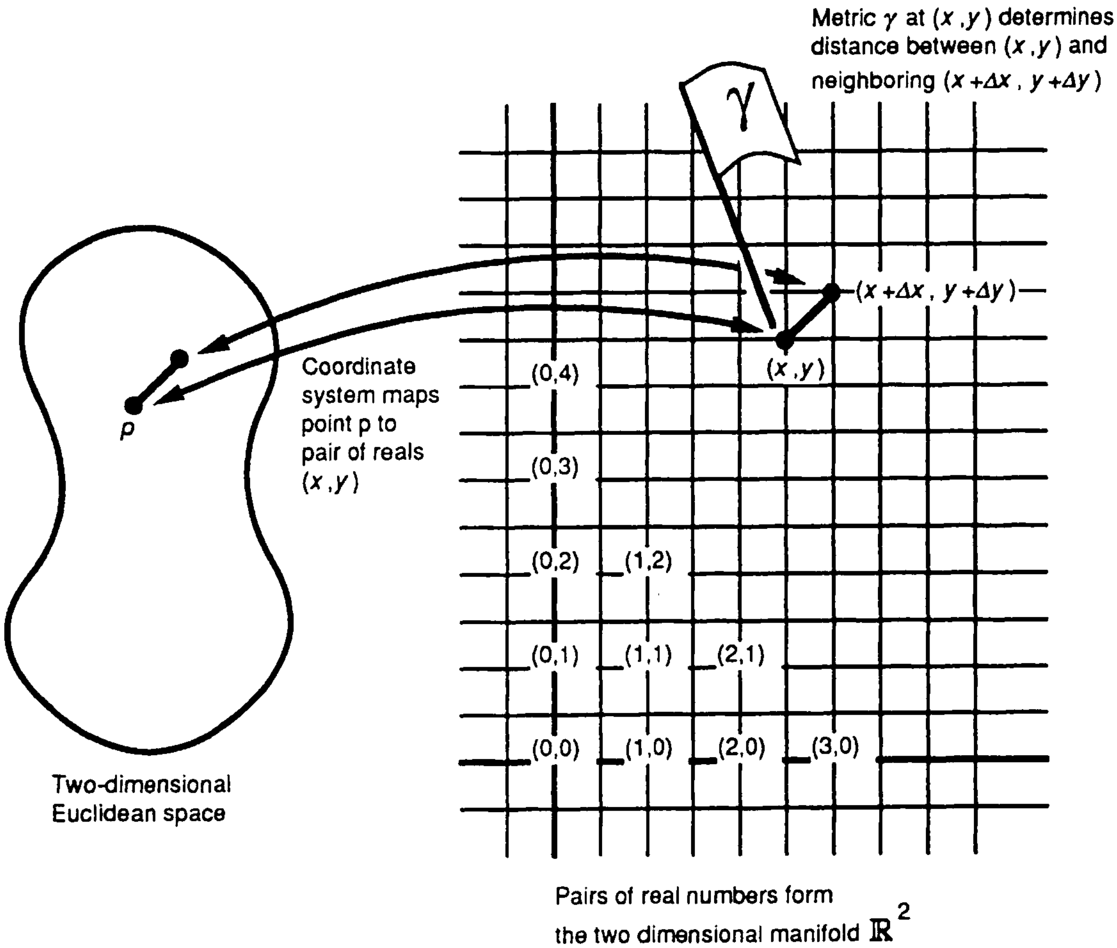


Figure 5.9 Model of a two-dimensional Euclidean space.

the manifold whose points are all the pairs of real numbers. Informally we picture the manifold $\mathbf{R}^2$ as the set of all pairs of real numbers laid out in a two-dimensional table that is without holes and that extends indefinitely. That this manifold can be coordinated with a physical Euclidean space reflects the fact that the space has all of its topological properties (see Figure 5.9). The theory is to be generally covariant. Therefore we allow any coordination between the physical space that is produced by a smooth transformation from the original. These transformations include all manner of translations, rotations, reflections, "stretchings" and "squeezings" that preserve smoothness of the coordinate system and the uniqueness of the labelling of the points.

Our theory cannot yet determine the distances between the points of the space. This information is provided by the geometric object γ which is the metric tensor of the space. This object is defined at every point of $\mathbf{R}^2$ and encodes the distances from that point to the points neighboring it. The metric tensor can be used to determine the length of curves in a Euclidean space by breaking up the curves into a sequence of small segments, determining the length of each segment and adding.

In sum, the models of the theory are pairs of the form

$$<\mathbf{R}^2, \gamma>$$

Since the theory is generally covariant, infinitely many coordinations will be allowed between the physical space and the manifold $\mathbf{R}^2$. Just as in the case of of the linear time theory, as we transform from one coordinate system to another, we may have to modify the scale factors forming γ to retain the invariance of the judgements of length which it hands down. Thus the model set of the theory will be infinitely large:

The distance $\Delta 1$ between a point with coordinates (x,y) and a neighboring point $(x+\Delta x, y+\Delta y)$ is given by the quadratic form

$$\Delta 1^2 = \gamma_{11}\Delta x^2 + \gamma_{12}\Delta x\Delta y + \gamma_{21}\Delta x\Delta y + \gamma_{22}\Delta y^2 \qquad (4)$$

where the coefficients γ_{12} and γ_{21} are equal. The matrix of the four values of these coefficients

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$$

represents the quantity γ in the relevant coordinate system. In certain special coordinate systems—the Cartesian coordinate systems—the coefficients reduce to an especially simple form ($\gamma_{11} = \gamma_{22} = 1$, $\gamma_{12} = \gamma_{21} = 0$) and (4) becomes

$$\Delta 1^2 = \Delta x^2 + \Delta y^2 \qquad (4')$$

which is a version of Pythagoras's theorem. A formulation of the theory of Euclidean space which uses only Cartesian coordinate systems is a standard formulation of the theory.

$$\langle \mathbf{R}^2,\gamma\rangle, \langle \mathbf{R}^2,\gamma'\rangle, \langle \mathbf{R}^2,\gamma''\rangle, \ldots$$

The quantities γ, γ', γ'' . . . transform into one another under transformation between different coordinations and jointly represent the one geometric object.

5.6 SYMMETRY PRINCIPLES

Symmetry principles provide a precise way of giving mathematical expression to important physical properties of space and time. In the theories of linear time and Euclidean space in Sections 5.4–5.5, symmetry principles express the homogeneity and isotropy of time and space. In the spacetime theories to follow, symmetry principles will also express the relativity principles of the theories.

The idea of symmetry used in analyzing these theories is no different in essence from the common notion of symmetry applied to everyday objects. One familiar type of symmetry is the bilateral symmetry exhibited (approximately) by the human form. To see the symmetry, imagine a transformation that switches the left- and right-hand sides of the body so that the left hand changes place with the right, the left foot with the right and so on. This transformation, a reflection about the central plane, is a symmetry of the human form since it leaves the form unchanged. Another type of symmetry is rotational symmetry exhibited, for example, by a cylinder. If we rotate the cylinder any number of degrees about its central axis, the rotated shape will coincide exactly with the unrotated shape (see Figure 5.10).

These examples illustrate the two essential elements of symmetry. First, one has a transformation, such as a reflection or rotation. Second, the transformation leaves something unchanged. The transformation is known as a *symmetry transformation* or, more briefly, a *symmetry* of that thing.

These same ideas can be applied to a Euclidean space as well. As a stepping-stone to this application, consider a pattern, such as we find on wallpaper. These patterns can exhibit symmetries. The pattern shown in Figure 5.11 exhibits a reflec-

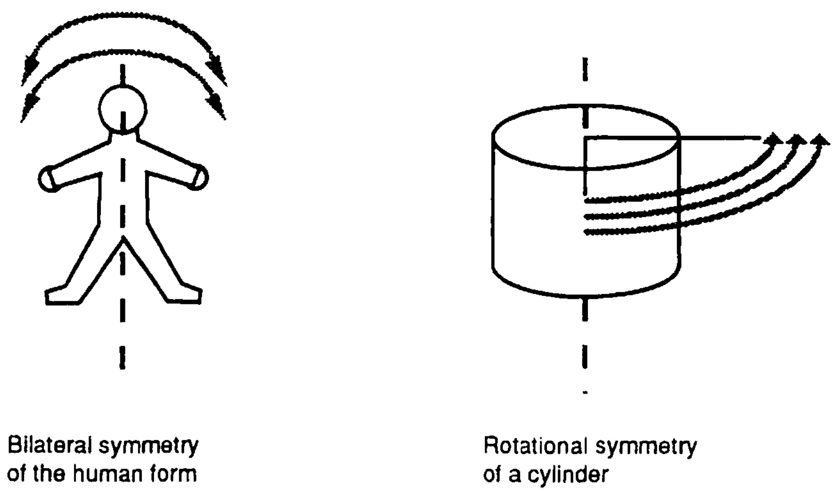


Figure 5.10 Symmetries of common objects.

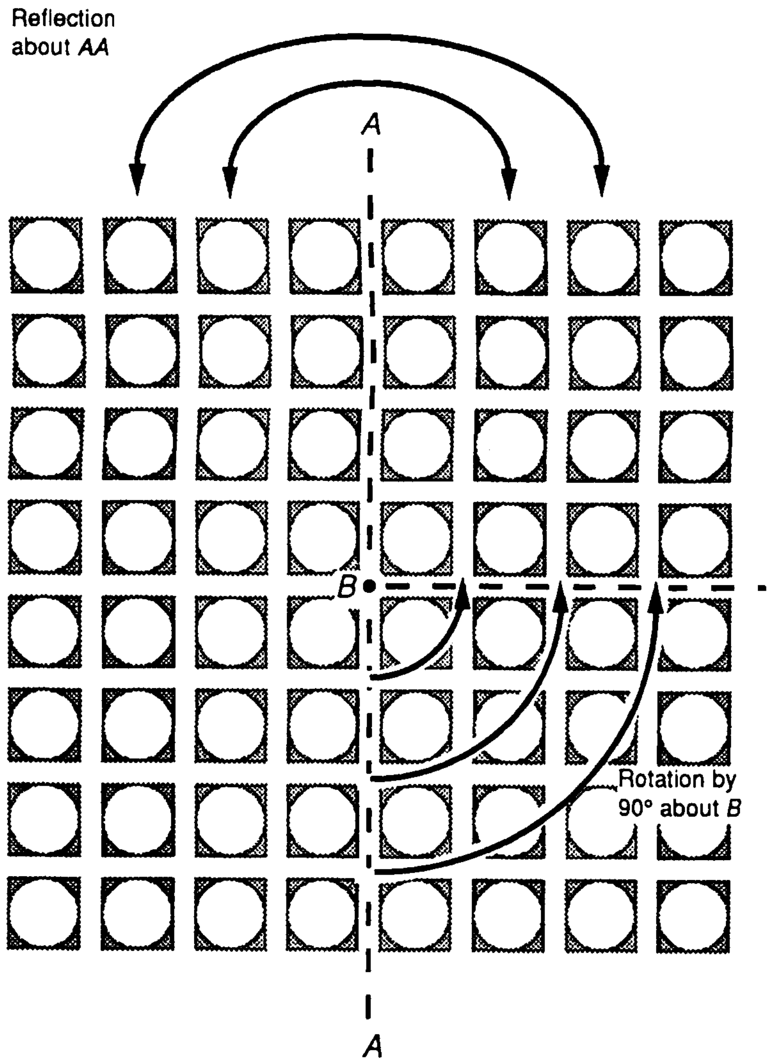


Figure 5.11 Symmetries of a decorative pattern.

tion symmetry since a reflection of the pattern about axis AA leaves the pattern unchanged. Similarly, the pattern exhibits a rotational symmetry. If the pattern is rotated by 90 degrees about the point B, then the pattern remains unchanged. In a Euclidean space $<\mathbf{R}^2,\gamma>$, the manifold $\mathbf{R}^2$ behaves like the paper and the metric γ is like the pattern painted on it. A transformation on this space that leaves the space unchanged is a symmetry transformation (or just symmetry) of the space. Three types of symmetry transformations are exhibited by this space as shown in Figure 5.12: a reflection about any axis, a rotation by any angle about any point, and a translation by any distance in any direction.

These symmetries of a Euclidean space express the space's homogeneity and isotropy. To say the space is homogeneous just means that every point and its geometry is exactly like every other point and its geometry. Thus if observers examine the geometry in the vicinity of one point of the space and then translate their viewpoint to any other point, then the geometry observed should remain unchanged.

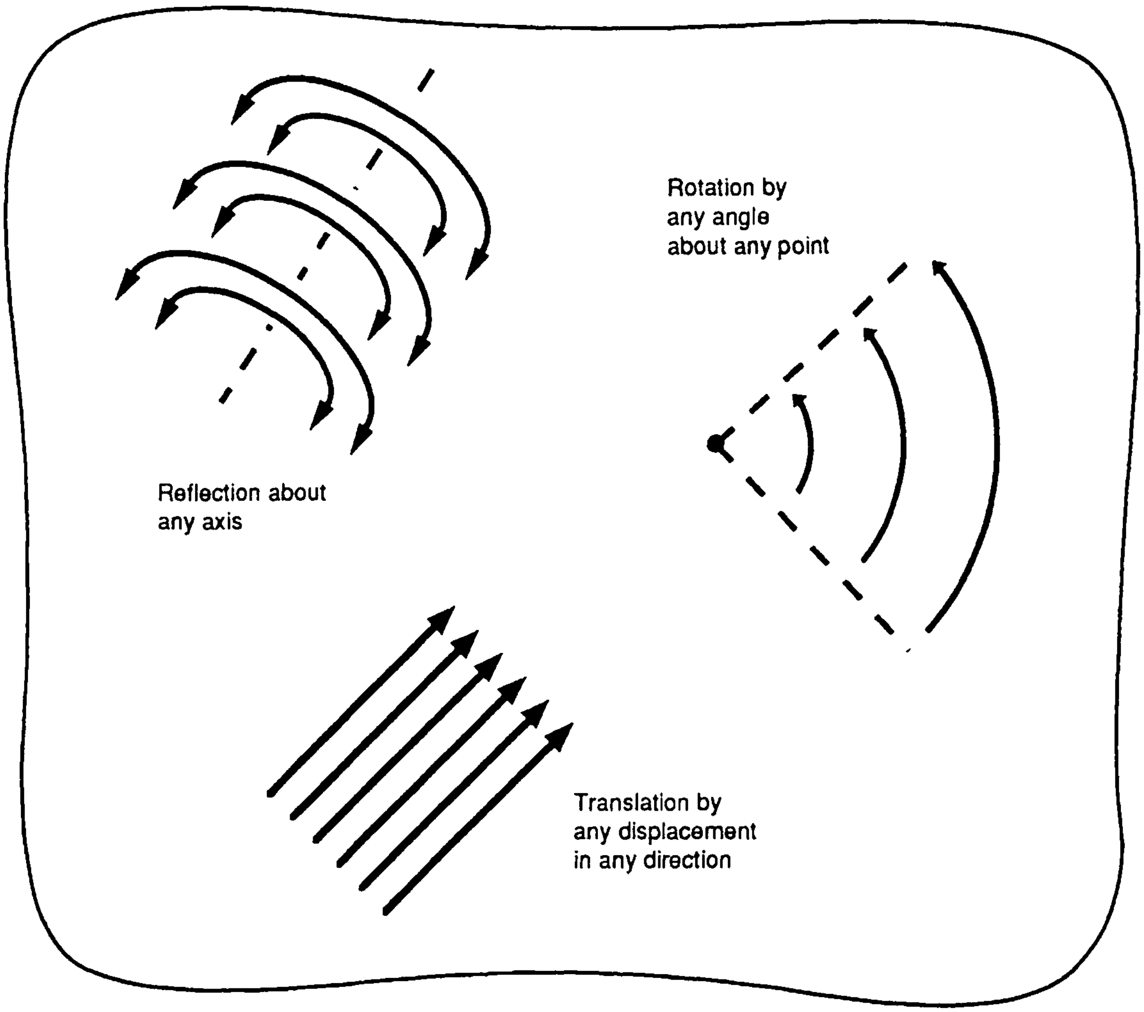


Figure 5.12 Symmetries of a two-dimensional Euclidean space.

But this merely says that any translation on the space leaves the space unchanged. That is, any translation is a symmetry of the space. Similarly, to say that the space is isotropic just means that every direction in the space is exactly like every other.

If the covariance group of a formulation of a theory of time, space or spacetime coincides with the group of its symmetry transformations, then that formulation is a *standard formulation* of the theory. A formulation of the theory of Euclidean space restricted to Cartesian coordinate systems is a standard formulation. Standard formulations tend to be simpler mathematically. However, they can be misleading since explicit mention of the geometric structures present tends to be simplified out of the formulation's equations. Thus the Euclidean metric γ is rarely mentioned in a standard formulation of the theory of a Euclidean space.

Thus if observers examine the geometry of the space as it lies in some direction at any point and then rotate their viewpoint by any number of degrees, then the geometry observed in the new direction should be the same. Again this merely says that any rotation about any point on the space leaves the space unchanged so that all such rotations are symmetries of the space.

5.7 NEWTONIAN SPACETIME

5.7.1 Transition to a Spacetime Formulation

In this section, let us develop a generally covariant, spacetime formulation of Newton's theory of space and time, modified to be compatible with the principle of relativity. To have such a formulation of the Newtonian theory for work in philosophy of space and time is very important, even though the new formulation is more complicated than the traditional one. Much philosophical interest exists in comparing the Newtonian theory with the theories of special and general relativity. The relativistic theories are presented most clearly in their generally covariant, spacetime formulations—general relativity necessarily so since no other formulation is known. For our comparisons to be reliable, we must carry them out on theories formulated in the same way. Otherwise our conclusions may well pertain not to true differences between the theories but only to differences between their methods of formulation. Section 5.10 discusses some of the damage that has been done by failing to use uniform formulations in such theory comparisons.

5.7.2 Formation

The Newtonian spacetime theory is produced by combining the theory of linear time with that of Euclidean geometry and just a little further structure. We begin with a Newtonian universe and take "snapshots" of its contents at all instants. These snapshots are simply three-dimensional Euclidean spaces (although for the figures we continue to suppress the third dimension and represent the space as a two-dimensional Euclidean space). Since each snapshot is taken at a different time, each of them can comprise an instant in the linear time theory. We construct the Newtonian four-dimensional spacetime by taking each of the three-dimensional Euclidean spaces and "stacking them up" in a linear continuum (see Figure 5.13). If we picture the spacetime as a deck of cards, then the geometry on each card (instantaneous snapshot) is given by a Euclidean metric γ. The temporal structure, as we proceed through the deck from card to card (instant to instant), is given by the temporal metric dT.

The deck of cards pictured shows us exactly where the theory as described so far is incomplete. Many ways are possible to stack up cards, as shown in Figure 5.13. Which is the right one? If we have points A, B, C . . . at rest in the space, then an acceptable stacking is one that places the points A, B, C . . . in each instant exactly on top of one another so that points at rest can be pictured as straight lines penetrating vertically through the stack. Moving points can also be represented as lines that penetrate obliquely through the stack. (To see this, imagine a point which moves from A to B to C as time proceeds from 0 to 1 to 2. It will be represented by a line that intersects A on the snapshot at time 0, B at 1 and C at 2.) In particular, we will represent points that move uniformly in a given direction—that is, move inertially—as *straight* lines penetrating the stack obliquely.

The stack of instants forms a four-dimensional manifold, each of whose points is an event, a point in space at a given time. Each instant is a three-dimensional surface in that manifold, technically a "hypersurface." These hypersurfaces are sets of simultaneous events, so they are called "*hypersurface of simultaneity.*" The lines representing moving and motionless points are their *worldlines*. The encode the entire history of each point's motion.

5.7.3 Principle of Relativity

The spacetime theory as described so far incorporates absolute rest. In assuming that there is only one correct way to stack the instantaneous snapshots, we have singled out the points A, B, C of Figure 5.13 as absolutely at rest. In section 5.1, we discussed the principle of relativity in terms of interpenetrating absolute and relative spaces. In the spacetime context, such spaces are represented by *frames of reference*. Consider the points of a relative space. Each point will be a worldline penetrating the stack of instants. The totality of points of the space will thus be represented by a dense bundle of worldlines penetrating the stack. If the space is an inertial space, then the corresponding bundle will be a bundle of straight lines as shown in Figure 5.14 and will be called an *inertial frame of reference*.

In effect Newton supposed that one of these inertial frames of reference was special and represented an absolute state of rest. Thus for him the only correct stacking of the surfaces of simultaneity would be one that aligned the points of this absolute frame. The principle of relativity requires that all inertial frames are to be equivalent so that all inertial states of motion are equivalent. Geometrically this amounts to saying that all directions in spacetime picked out by inertial frames are equivalent. Thus if we consider two inertial frames, such as in Figure 5.14, we should not think of either as having properties different from the other. Unfortunately, because of the limitations of drawing pictures of inertial frames, one frame is drawn as penetrating the stack vertically and the other obliquely. This difference is not reflected in the actual geometric structure of a Newtonian spacetime.

The situation is closely analogous to the isotropy of a Euclidean space. All directions in such a space are equivalent. However if we draw a picture of these directions, such as in Figure 5.15, one points up the page in the 0-degree direction and another across it in the 90-degree direction. Since a Euclidean space admits rotations as a symmetry, we can erase any suggestion that a given direction in the space is preferred by rotating the space so that the given direction is at the 0-degree position and noting that the space is unchanged.

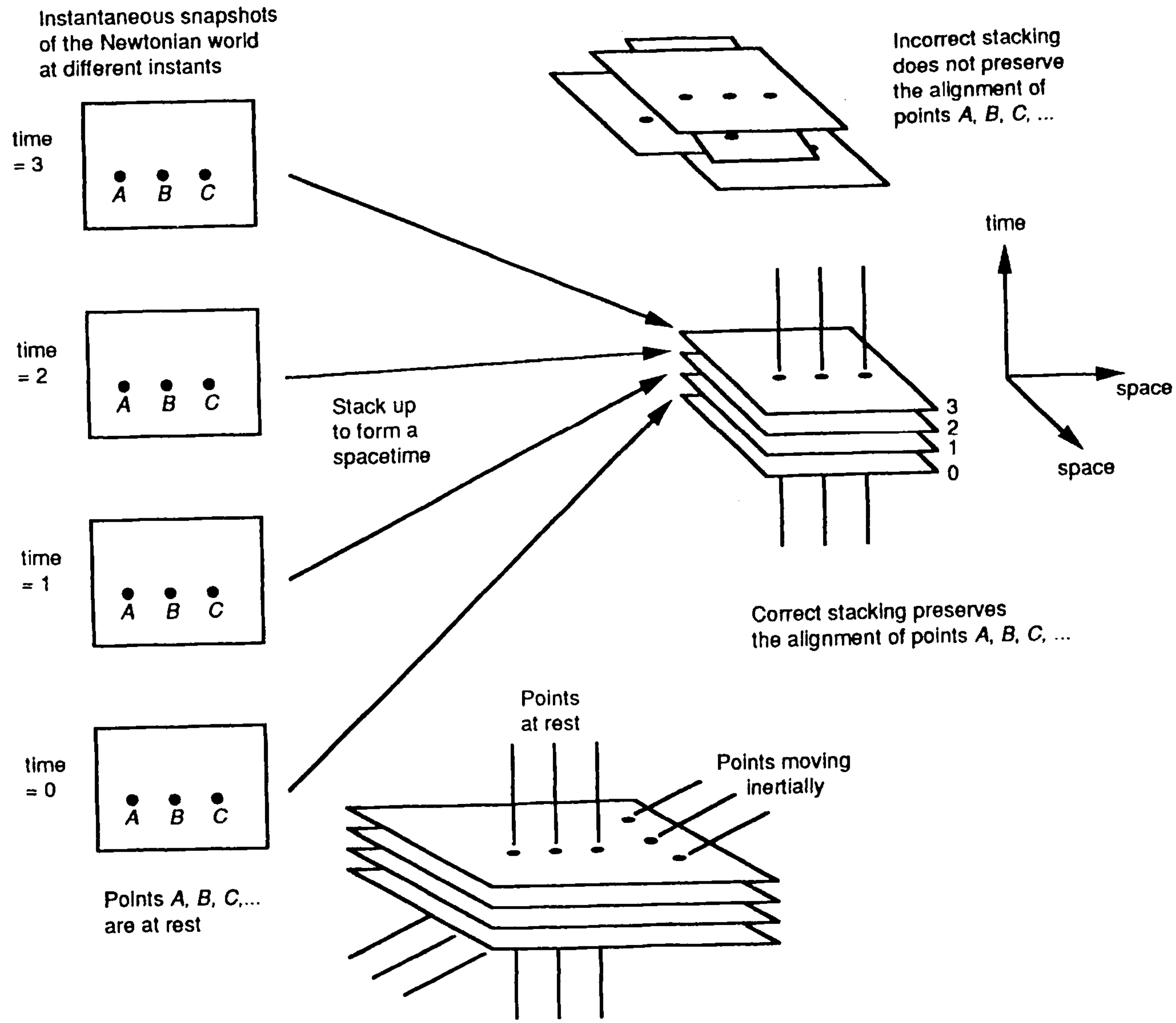


Figure 5.13 The formation of a Newtonian spacetime.

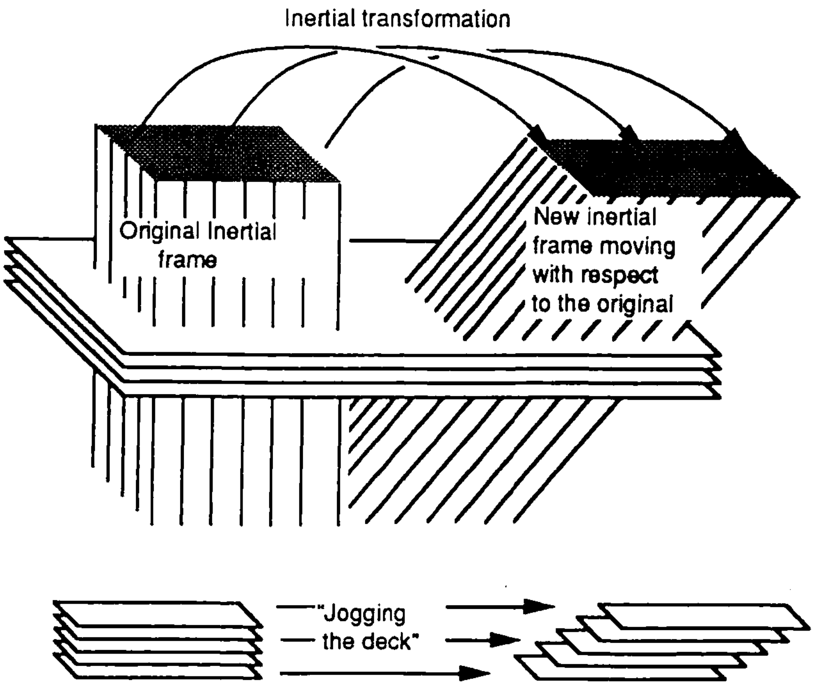


Figure 5.14 Inertial transformation in a Newtonian spacetime.

Similarly, the rules for stacking the deck of hypersurfaces of simultaneity must allow us to restack the deck so that it is aligned according to any inertial frame of reference. This means that any frame can be transformed to the zero velocity state. Let us call the transformation that maps inertial frames into inertial frames, shown in Figure 5.14, an "*inertial transformation.*" Figuratively it corresponds to realigning the hypersurfaces of simultaneity in a manner akin to jogging a deck of cards. What we have concluded is that an inertial transformation cannot change the spacetime in the same way that a rotation does not change a Euclidean space, so that the picking out of any inertial frame as uniquely at rest is a purely arbitrary stipulation. That is, the principle of relativity is a symmetry principle:

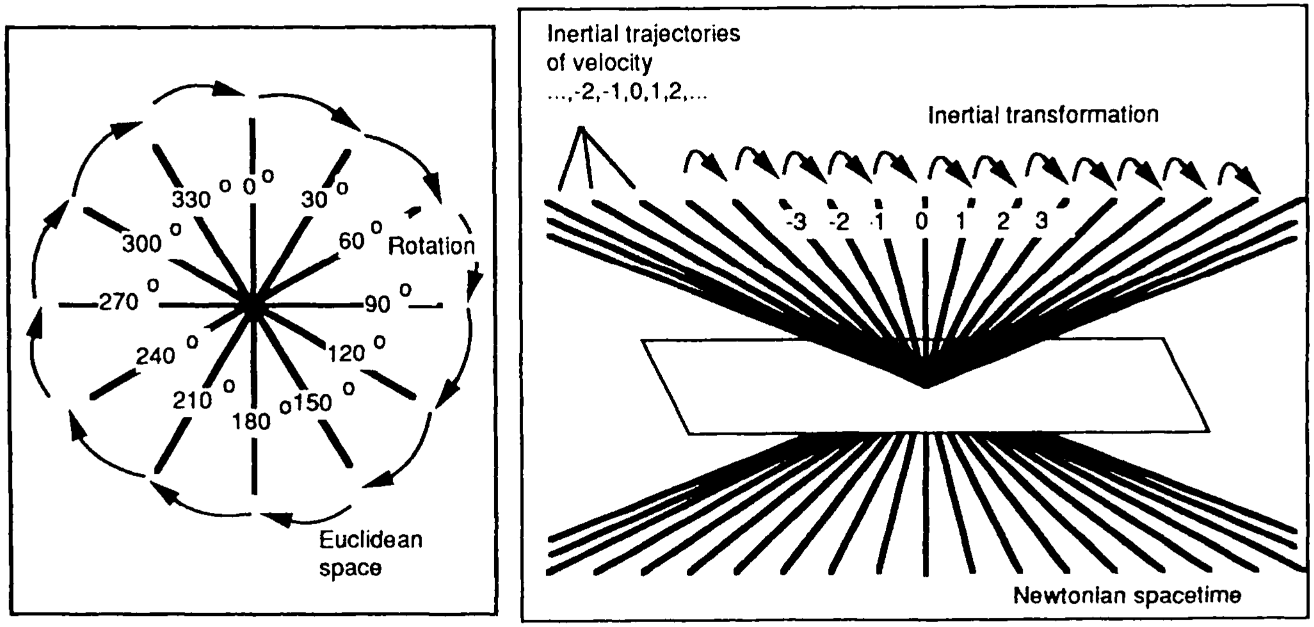
> *Principle of relativity in a Newtonian Spacetime:* An inertial transformation is a symmetry of a Newtonian spacetime; it leaves the spacetime unchanged.

5.7.4 Models of a Newtonian Spacetime

To summarize, the models of a Newtonian spacetime are quadruples

$$<M, dT, h, \nabla>,$$

where M is a four-dimensional manifold each of whose points represents an event. This manifold is sliced into instants, that is, hypersurfaces of simultaneous events. The measurable time elapsed as we move from instant to instants is given by the temporal metric dT. Each of the hypersurfaces of simultaneity is a Euclidean space with its own Euclidean metric γ; the structure h combines all of them into a single



No preferred direction in Euclidean space since any line of figure can be rotated to 0° position.

No preferred inertial velocity in Newtonian spacetime since any inertial trajectory can be transformed to 0 velocity.

Figure 5.15 Symmetries of Euclidean space and Newtonian spacetime.

geometric object. Finally we need a structure which will dictate the allowed ways of stacking instants. That structure is the *affine structure* ∇ of the spacetime. The affine structure specifies which of all the curves in the four-dimensional manifold M are the straight lines. (Notice that neither structure introduced so far—neither the Euclidean metric of each hypersurface of simultaneity nor the temporal metric—gives us any way of determining which are the straight lines that penetrate through the hypersurfaces.) We require that the instants be stacked in such a way that the trajectories of inertially moving points coincide with the straight lines of the manifold's affine structure. This rule will be compatible with the principle of relativity, if we require that the affine structure ∇, as well as temporal and spatial metrics dT and h, admit inertial transformations as symmetry transformations.

We recover a standard formulation of Newtonian spacetime theory by adopting standard coordinates T from the linear time theory and X, Y and Z from Euclidean geometry and combine them to form a coordination between the Newtonian spacetime and $\mathbf{R}^4$. The straight lines of the affine structure ∇ are now just what you would expect: the set of all lines given by the linear relations between the coordinates including $T = aX = bY = cZ$, for all real values a, b and c. A typical inertial frame is given by the set of all such straight lines parallel to the T axis. An inertial transformation that transforms this frame to a frame moving at velocity V in the X direction is given by

$$T' = T, \qquad X' = X - VT, \qquad Y' = Y, \qquad Z' = Z.$$

5.8 SPECIAL RELATIVITY

5.8.1 Relativity of Simultaneity

Einstein developed his special theory of relativity in 1905 axiomatically as the consequences of two postulates: the principle of relativity and what we now call the *light postulate*. The latter postulate asserts that the velocity of light has the same constant value (c = 300,000 km/sec) in all inertial spaces. On first acquaintance, it seems that no theory free of logical contradiction could be based on these postulates. How could the velocity of light remain the same in all inertial spaces? Surely if we transform to inertial spaces moving successively faster in the direction of a light ray, the light ray's velocity must be diminished as we catch up with it until it is finally brought to a standstill. The light postulate asserts that we can never catch the light ray. No matter how fast we go in chasing it, it always moves away from us at the same speed, 300,000 km/sec. What Einstein realized was that this state of affairs was possible if we were prepared to forgo some commonly assumed properties of space and time. One of the most important concerned simultaneity. In the Newtonian theory it had been assumed that two events either were or were not simultaneous. In special relativity, things ceased to be so simple.

Consider Einstein's standard simultaneity relation defined in Section 5.3. Assume that clocks A and B of Figures 5.2 and 5.3 have been synchronized by Einstein's light signaling procedure so that they are in standard synchrony (at least according to an observer at rest with respect to them). If we now transform our viewpoint to an inertial space in which clocks A and B are moving together in the direction from A to B, we no longer agree that the two clocks are in standard synchrony. In the new inertial space, the light signal will have to traverse a greater distance on its outward journey than on its return journey. For on the outward journey it must catch a B-clock that flees from it, whereas on the return journey the A-clock rushes forward to meet it. If the light postulate is correct and the speed of light has the same constant value in the new inertial space in both directions, then the outward journey must take more time than the return journey, so that the event of the reflection of the signal at B cannot happen midway between its emission and return at A—at least according to an observer in the new inertial space. That is, the clocks cannot be in standard synchrony in the new inertial space.

Thus in special relativity judgements of whether two clocks are in standard synchrony and, therefore, whether two events are simultaneous depend on the choice of inertial space to which the judgements are referred. This result is known as the *relativity of simultaneity*. It should not be confused with the *conventionality* of simultaneity discussed in Section 5.3. The relativity of simultaneity applies even after a particular definition of simultaneity has been chosen, such as standard $\epsilon = 1/2$ simultaneity above,and arises when we change inertial spaces. The conventionality of simultaneity arises within a single inertial space.

5.8.2 Minkowski Spacetimes and the Lorentz Transformation

The four-dimensional spacetime formulation of special relativity was discovered by Hermann Minkowski in 1907. Its spacetimes are called Minkowski spacetimes in his honor. A Minkowski spacetime is much like a Newtonian spacetime. Both are based on four-dimensional manifolds of events. Moving points in each are curves, and points moving inertially are straight lines, so that inertial frames of reference are still bundles of parallel straight lines. However, the most prominent landmark of a Newtonian spacetime, its unique divisibility into hypersurfaces of simultaneity, is not present in a Minkowski spacetime. For the relativity of simultaneity entails that each inertial frame defines a different slicing of the spacetime into hypersurfaces of simultaneous events. A hypersurface of simultaneity of a given inertial frame of reference is said to be *orthogonal* to the curves of the frame.

The transformation between inertial frames of reference in special relativity is called a *Lorentz transformation*. The relativity of simultaneity makes it more complicated than in the Newtonian case shown in Figure 5.14. For in the Lorentz transformation, the slicing of the spacetime into hypersurfaces of simultaneity must be

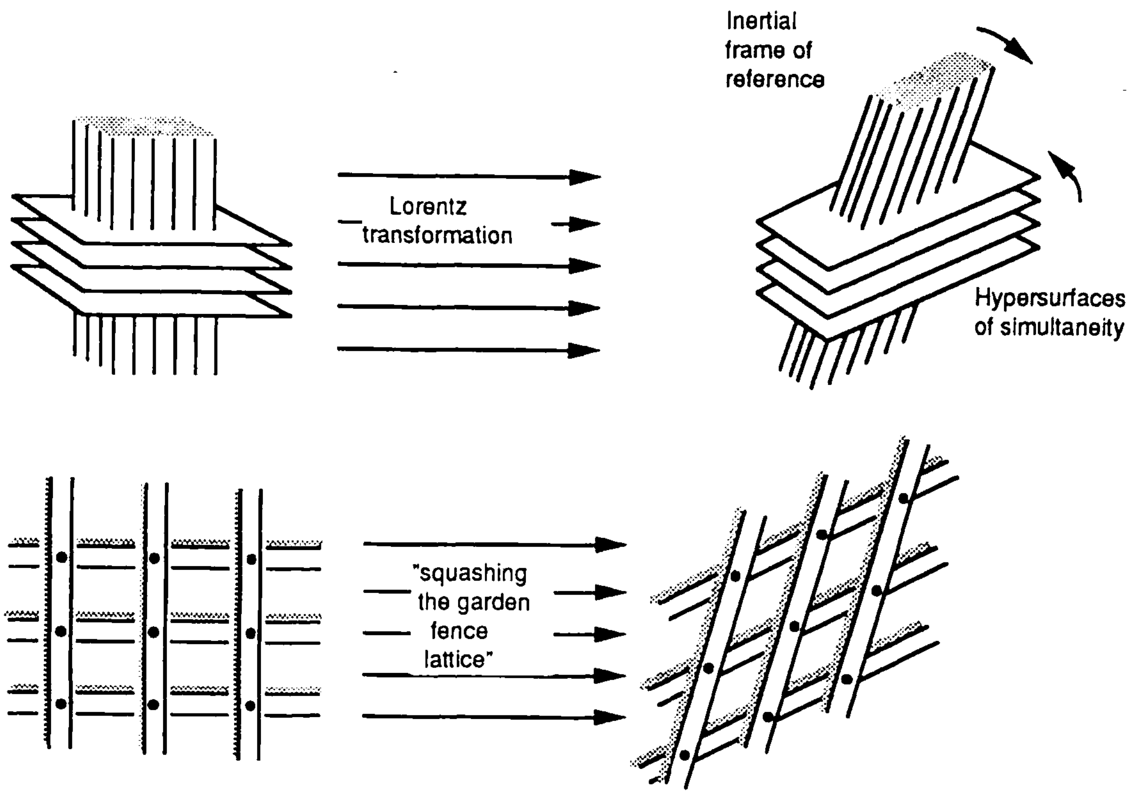


Figure 5.16 The Lorentz transformation.

adjusted to the new frame so that the hypersurfaces of the slicing remain orthogonal to it. The transformation is pictured in Figure 5.16. Where the inertial transformation of a Newtonian spacetime is mechanically akin to "jogging a deck of cards," the Lorentz transformation is mechanically akin to "squashing the garden fence lattice."

Finally we note that the Lorentz transformation is a symmetry of the Minkowski spacetime so that the principle of relativity holds just as it does in the Newtonian spacetime of the previous section.

> In a standard formulation of special relativity, the standard coordinates X, Y, Z and T correspond to space and time measurements made by instruments at rest in the frame whose worldlines are the T curves. The Lorentz transformation, which transforms this frame to one moving at velocity V in the X direction, is given by
>
> $$T' = \beta(T - VX/c^2) \qquad X' = \beta(X - VT) \qquad Y' = Y \qquad Z' = Z$$
>
> where
>
> $$\beta = 1/\sqrt{1 - V^2/c^2}.$$

5.8.3 Light Cone and Causal Structures of a Minkowski Spacetime

An infinite number of curves pass through any given event of a Minkowski spacetime. The light cone structure of the Minkowski spacetime at that event is simply the division of the curves *at that event* into three classes: those that represent

1. points moving at velocity c, the velocity of light (''light-like'');
2. points moving at velocity less than c (''time-like'');
3. points moving at velocity greater than c (''space-like'').

The name ''light cone'' arises from the fact that the light-like curves form a cone through the event as shown in Figure 5.17. The time-like curves all fall within the cone and the space-like curves outside the cone. The light cone structure of the entire Minkowski spacetime is the specification *at every event* of the above three-way division.

Time-like curves can be the worldlines of massive particles. Light-like curves can be worldlines of light signals. The usual assumption in special relativity is that no causal process such as a particle or signal can travel faster than light so that space-like curves *cannot* be the worldlines of any particle or signal. Under this assumption, the light cone structure takes on special significance for the philosophy of space and time, for it is equivalent to the causal structure of the spacetime. More precisely, if we know the light cone structure of the spacetime then we can construct an exhaustive catalog of which pairs of events can causally interact with one another in the spacetime. We do this by finding all pairs of events which could be connected by the trajectory of a particle or signal, that is, by a curve that is everywhere time-like or light-like (see Figure 5.18). The resulting catalog is the causal structure of the spacetime. Conversely, if we know this catalog, then we can reconstruct the light cone structure.

5.8.4 The Minkowski Metric

As a *spacetime* theory, the Newtonian theory is rather complicated. It requires three distinct structures to be specified: dT, h and ∇. As a spacetime theory, special relativity is far simpler. The functions of the three Newtonian structures is performed by just one, the Minkowski metric η. Thus models of special relativity are of the form

$$<M,\eta>$$

where M is a four-dimensional manifold and η a Minkowski metric. The properties of η are very similar to those of a Euclidean metric γ (see the following box) since η also assigns lengths—called ''intervals''—to curves. The metric η picks out which are the time-like, space-like and light-like curves by the intervals it assigns to them. It assigns a zero interval to light-like curves. It assigns a positive interval to time-like

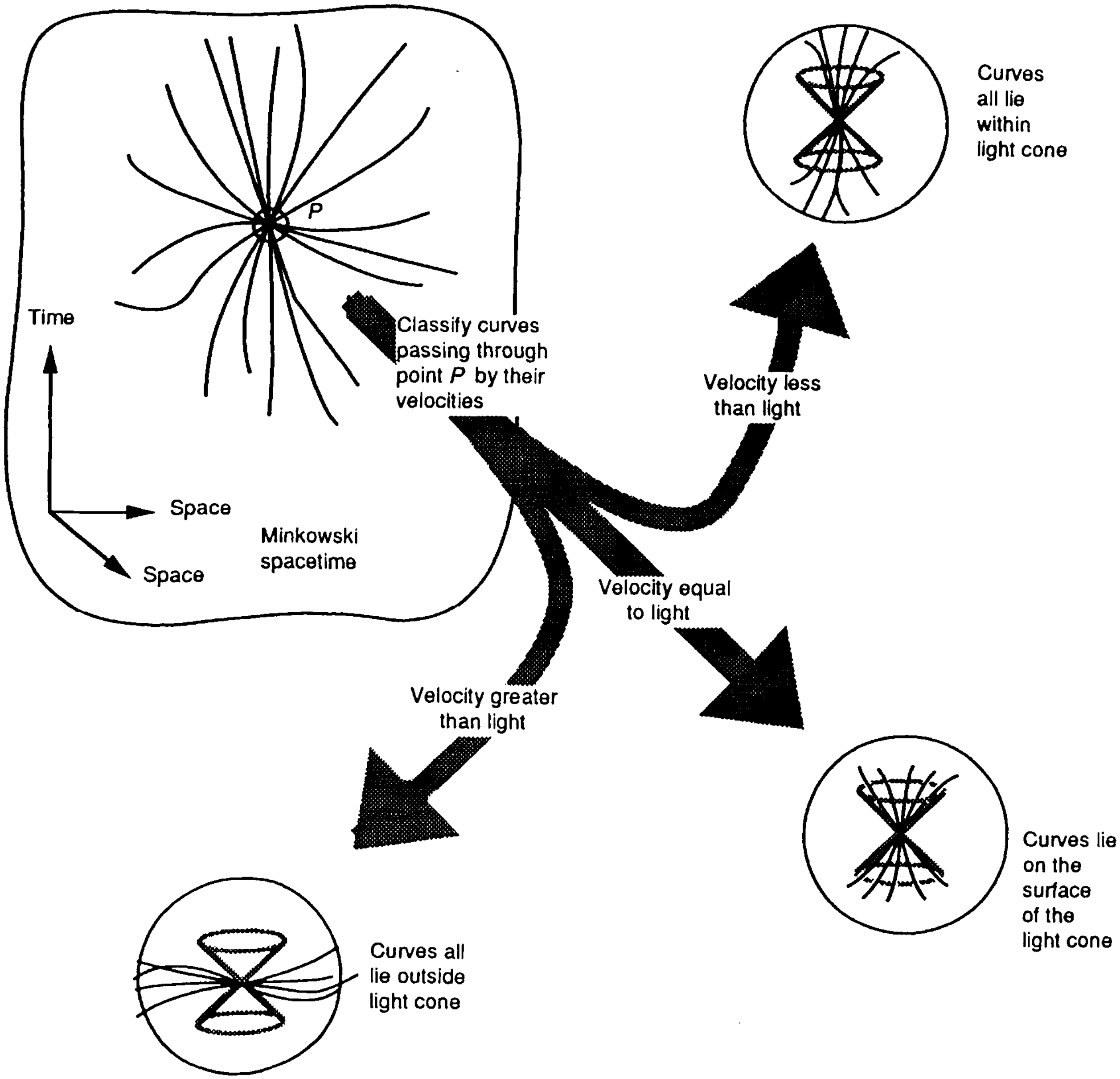


Figure 5.17 Light cone structure at an event in a Minkowski spacetime.

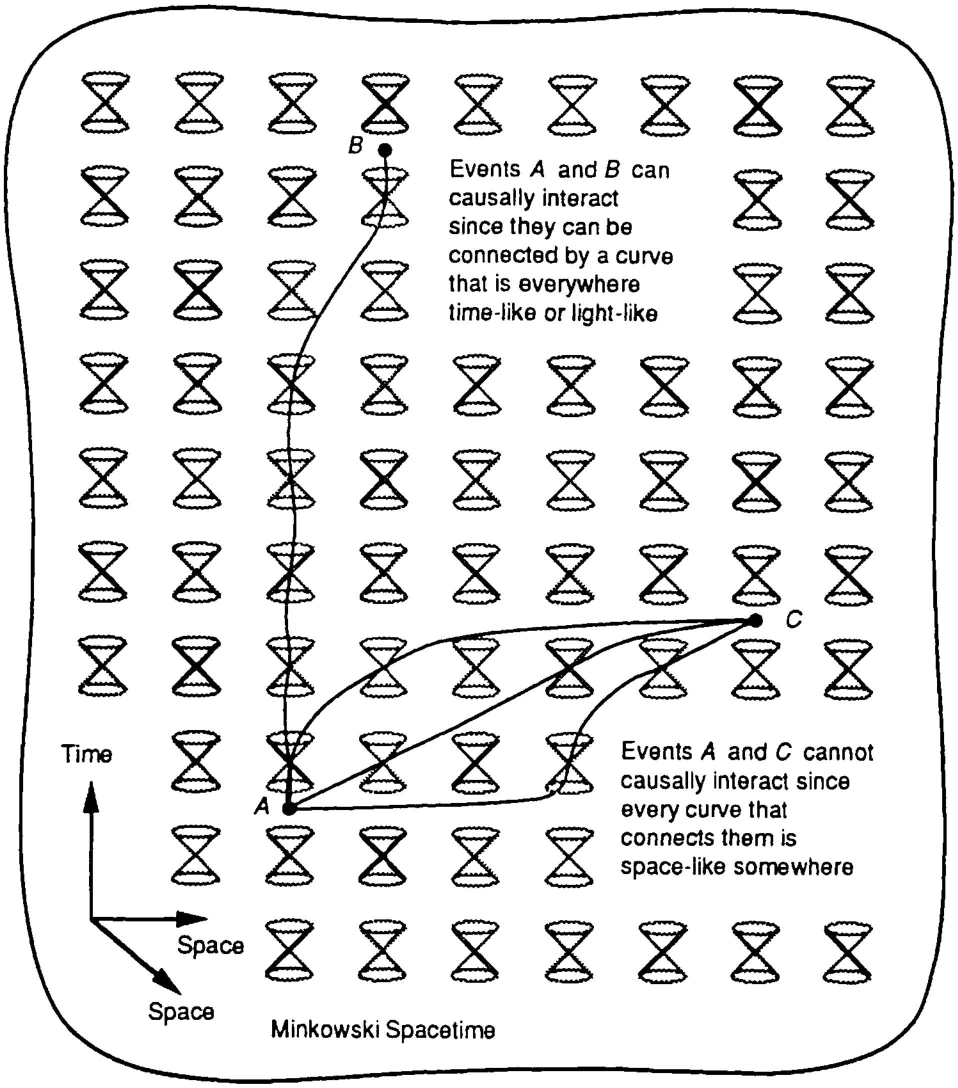


Figure 5.18 Light cone structure determines causal structure.

In a standard coordinate system X, Y, Z, T of special relativity, the Minkowski metric is associated with the differential form

$$\Delta s^2 = \Delta T^2 - \Delta X^2 - \Delta Y^2 - \Delta Z^2 \qquad (5')$$

which is fully analogous to the differential form (4′) of the Euclidean theory. The differences between the metrics of the two theories derive entirely from the differences in sign between forms (4′) and (5′). The transition to the generally covariant formulation introduces four arbitrary spacetime coordinates, x^0, x^1, x^2, x^3 and the equation (5′) generalizes to

$$\Delta s^2 = \eta_{00}(\Delta x^0)^2 + \eta_{01}\Delta x^0 \Delta x^1 + \ldots + \eta_{23}\Delta x^2 \Delta x^3 + \eta_{33}(\Delta x^3)^2 \qquad (5)$$

which is analogous to (4). The explicit representation of the metric η is the symmetric matrix of coefficients η_{ik}, where $i,k = 0, 1, 2, 3$.

curves. This interval is the time elapsed as measured by a clock moving with the particle represented by the curve. The Minkowski metric assigns an imaginary interval to space-like curves. The absolute value of this interval is the spatial length of the curve should the curve lie fully in a hypersurface of simultaneity.

5.9 GENERAL RELATIVITY

5.9.1 Physical Foundations

General relativity is Einstein's relativistic gravitation theory and is a modification of special relativity that incorporates gravitation. It was completed by him in 1915 and is probably his greatest contribution to physics. The novelty of the theory is the way that gravitation is treated. Prior to general relativity, it was customary to think of a gravitational field as a distinct entity that could be added to a spacetime. Thus gravitation-free spaces were possible. In general relativity, the gravitational field is combined with the same structure that determines lengths and times so that a gravitation-free space is no longer possible.

The chain of ideas that led Einstein to general relativity began in 1907 when he was struck by a remarkable property of gravitation known since the time of Galileo. When a gravitational field deflects the motion of a body, the amount of deflection is independent of the nature of the body and, in particular, the mass of the body. This property is a very special property of gravitational fields and is not shared, for example, by electric fields. If the motion of a charged body is deflected by an electric field, then the greater the charge on the body, the larger the deflection. It was as though the trajectories of bodies falling in a gravitational field were already laid out in spacetime and any falling body would have to follow them, whatever its mass. Now a Minkowski spacetime just happens to have trajectories with exactly this property. These are the trajectories of inertially moving points, the straight time-like worldlines defined by the Minkowski metric. Any body moving inertially in a Minkowski spacetime follows these trajectories in a way that is independent of the mass of the body. Since these trajectories have exactly the unique, characteristic property of gravitation, Einstein was drawn to conjecture that a Minkowski spacetime was actually already a special case of a spacetime with a gravitational field and that spacetimes with more general gravitational fields could be constructed not by adding further structures to the spacetime but by modifying what was already there.

5.9.2 Principle of Equivalence

This conjecture was formulated and justified in a vivid manner in a thought experiment. Einstein imagined a physicist enclosed in a box in the supposedly gravitation-free space of special relativity. He then imagined that the box was accelerated uniformly in some direction. All free objects in the box would fall to one side with the same acceleration. The observing physicist, Einstein argued, could explain this phenomenon in two equally good ways. He could say that the box was accelerated. Alternately, because of the special property of gravity, he could say that the box was unaccelerated but that a homogeneous gravitational field was acting on the box. Einstein's "principle of equivalence" asserts that the two states of affairs—uniform acceleration in a gravitation-free space and a homogeneous gravitational field—are fully equivalent or, in our words, exactly the same state of affairs. Reduced to its briefest form, the thought experiment shows us that supposedly gravitation-free special relativity already incorporates gravitation—to see that gravitation is already there, transform to a uniformly accelerated space to make a homogeneous gravitation field manifest.

5.9.3 Generalizing Special Relativity

What characterizes the gravitational fields of special relativity is the following property: If two test bodies have initial velocities identical in magnitude and direction, they will continue to move so that the distance between them remains the same (see Figure 5.19). We are interested in more general gravitational fields such as those produced by the earth. In these more general cases, the distance between the above two bodies would not remain constant but would converge or diverge as the bodies fell. To construct general relativity, Einstein replaced the Minkowski metric η of special relativity with a more general metric g which would allow this convergence or divergence. In the new theory, unrestrained particles still follow the straight time-like curves of the spacetime, just as they did in a Minkowski spacetime. However the "straight" lines defined by the new more general metric g no longer behave in the way that we expect straight lines to behave. For example, two "straight" lines that are initially parallel need not remain at a constant distance from one another as

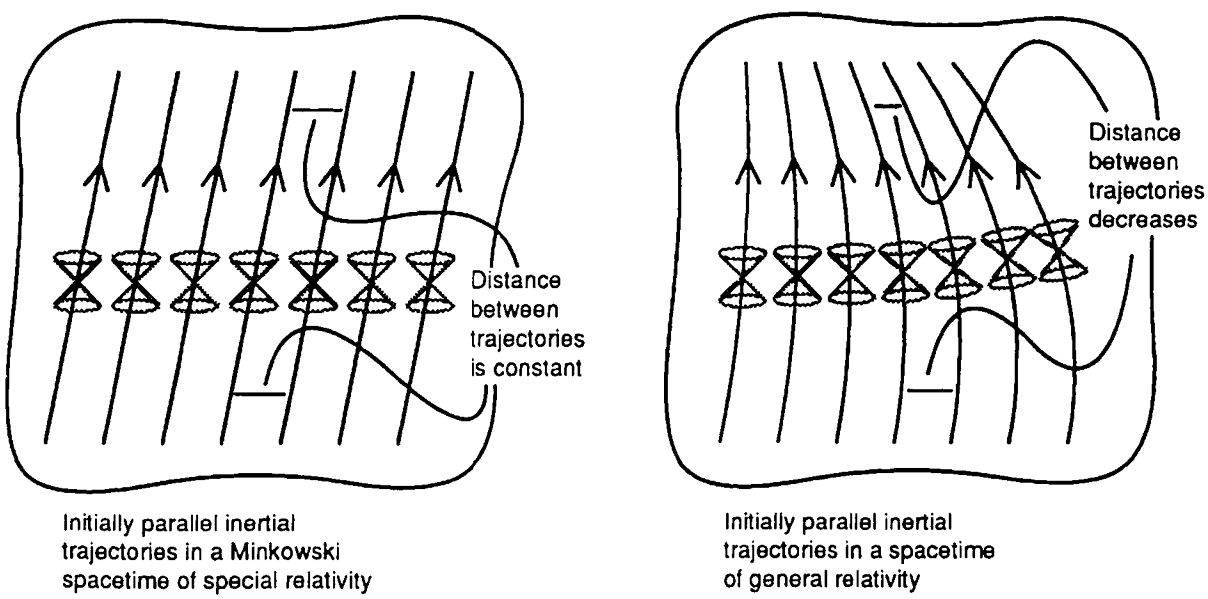


Figure 5.19 Inertial trajectories in special and general relativity.

we proceed along them (see Figure 5.19). Such results are typical in the geometry of curved surfaces, such as the surface of a sphere, and the mathematical techniques used in general relativity were originally developed in the context of problems of curved surfaces. As a result, talk of "curvature" is common and we routinely distinguish the "flat" spacetime of special relativity from the "curved" spacetimes of general relativity.

In sum, the models of general relativity have the form

$$<M,g>$$

where M is a four-dimensional manifold and g is a generalization of the Minkowski metric η. Since every distinct distribution of masses in the universe produces a distinct gravitational field, there will be very many different models in the theory. In particular, a nonuniform matter distribution will produce a nonuniform gravitational field. As a result, the models of general relativity will, in general, have no nontrivial symmetries, so that we cannot formulate relativity principles of the type seen in the flat Newtonian spacetime theory and special relativity.

Part III: Applications

5.10 CONFUSIONS OVER COVARIANCE

Misunderstandings of the significance of the covariance group of a theory have been responsible for more than their fair share of confusions in philosophy of space and time. Let us review two important examples.

5.10.1 The Generalization of the Principle of Relativity

One of Einstein's best known claims for his general theory of relativity is that it extends the principle of special relativity to accelerated motion. We noted in the previous section that the spacetimes of general relativity admit no nontrivial symmetries in general, so that we cannot formulate a relativity principle of the type formulated in Newtonian theory or special relativity. Thus Einstein's claim has proved increasingly difficult to defend and its defense has required stratagems of increasing complexity. (Friedman 1983 makes the case against the claim especially clear.) The simplest and most common argument for the claim is not a good one. It merely notes that general relativity is a generally covariant theory. However, general covariance by itself cannot sustain the claimed generalization of the principle of relativity since every spacetime theory we have examined in this chapter has been given generally covariant formulation. They cannot all satisfy a generalized principle of relativity!

The illusion that general covariance and an extension of the principle of rela-