Generalization of the Special Theory of Relativity

The general laws of nature are to be expressed by equations which hold good for all systems of coordinates, that is, are co-variant with respect to any substitutions whatever (generally co-variant).

A. Einstein

2.1 The Idea of Covariance

Minkowski made the important discovery that transforming from one inertial frame to another moving with relative velocity v corresponds (1) to the rotation of axes in a four-dimensional space-time coordinate system. The requirements of special relativity are indeed met in a most elegant fashion by writing the laws of physics as relations among four-dimensional vectors. While this procedure is not essential, it does add elements of beauty and simplicity.

We employ coordinates labeled by superscripts. In the Minkowski space of special relativity we have the coordinates x, y, z, and ct, which we will denote x^1, x^2, x^3, and x^0. The important quantities are the relations between events. An event has no extension in space or time. It is a point in a four-dimensional space. The interval between two events a and b is denoted by the symbol s_{ab} and defined by

$$s_{ab}{}^2 = (x_a{}^0 - x_b{}^0)^2 - (x_a{}^1 - x_b{}^1)^2 - (x_a{}^2 - x_b{}^2)^2 - (x_a{}^3 - x_b{}^3)^2 \quad (2.1)$$

If we transform from one Minkowski system of coordinates to another, the interval between the same events remains invariant and in terms of new coordinates x'^1, x'^2, x'^3, x'^0 is given by

$$(s_{ab})^2 = (x'^0{}_a - x'^0{}_b)^2 - (x'^1{}_a - x'^1{}_b)^2 - (x'^2{}_a - x'^2{}_b)^2 - (x'^3{}_a - x'^3{}_b)^2 \quad (2.2)$$

While the magnitudes of some quantities such as s_{ab} are invariant, others such as $x_a^1 - x_b^1$ do change under these coordinate transformations, which constitute a group. The form of the relation for s_{ab} does not change, and such a relation is said to be covariant under the group of coordinate transformations (rotation of axes in Minkowski space). For two events separated by a differential interval we can write

$$ds^2 = dx^{0^2} - dx^{1^2} - dx^{2^2} - dx^{3^2} \tag{2.3}$$

2.2 The Metric Tensor

We shall employ a great many sums and shall follow the summation convention that an index which is repeated represents a sum. This enables us to write (2.3) in the form

$$-ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu \tag{2.4}$$

The elements of $g_{\mu\nu}$ are represented by

$$g_{\mu\nu} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{bmatrix} \tag{2.5}$$

$g_{\mu\nu}$ is called the metric tensor and the expression (2.5), which retains its form under Lorentz transformations, is called a Lorentz metric.

In curvilinear coordinates, the metric tensor assumes a different form, but a coordinate transformation can transform it to the form (2.5), throughout the space, provided Euclidean geometry is valid there.

2.3 The Metric Tensor in Curved Spaces and Accelerated Frames

The curved two-dimensional space on the surface of a sphere is described by the squared line element

$$g_{\mu\nu}\, dx^\mu\, dx^\nu = r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2;$$

$$g_{\mu\nu} = \begin{vmatrix} r^2 & 0 \\ 0 & r^2 \sin^2\theta \end{vmatrix} \tag{2.6}$$

Triangles composed of arcs of great circles have the sum of their angles greater than π and less than 3π. This non-Euclidean, curved character is implied by the metric (2.6), through relations involving the derivatives of $g_{\mu\nu}$. No coordinate transformation can reduce (2.6) everywhere to a diagonal form in which the diagonal elements are one. The concept of curvature will be discussed in more detail later.

Suppose now that we have a plane triangle in an inertial frame. If we measure from a different inertial frame, the shape of the triangle may change, but the sum of its angles remains π and the geometry will not have changed.

Consider a set of points which lie on a circle in an inertial frame. If we observe from a rotating frame of reference with center coinciding with the center of the circle and survey the same points, we find that the ratio of circumference to diameter now depends on our operational procedure for measuring length. To avoid, for the moment, the issue of clock synchronization within a rotating system, we assume that length is measured by bringing pairs of points into coincidence using the clocks of the inertial frame to define simultaneity. The ratio of circumference to radius using measuring rods at rest within the rotating system will exceed π. This follows because a rigid measuring rod would measure the same diameter as before, but when laid along the circumference it would be foreshortened by the Lorentz contraction. The metric tensor in the rotating frame now has to describe a non-Euclidean, curved space.

In the inertial frame we may set up a system of fixed synchronized clocks throughout the space. In the rotating frame clocks at different radii will have different time measures. Identical clocks at different radii are similar to identical clocks in different inertial frames. They cannot be synchronized. Suppose that coordinate time is measured by pulses emitted by a clock at the center of the rotating frame. A clock at each point in the rotating frame ticks intervals for an observer at that point. The relation between these intervals and the coordinate time differences corresponding to receipt of successive light pulses will be described by the g_{00} component of $g_{\mu\nu}$. This will now be a function of radius, and the metric (2.5) of special relativity does not apply.‡

The equivalence of a gravitational field to an accelerated§ frame then implies that the special theory of relativity cannot be valid in an extended region (2) where gravitational fields are present. A curved-space metric is needed. To consider another example, suppose we make up a triangle in an inertial frame with sides which are light rays. The sum of the angles is π. If this is repeated in a gravitational field the sides of the triangle will become curved because of the action of gravitation on the energy of the photons. The sum of the angles of the triangle will now differ from π. In a covariant theory this is described by saying that the paths of light rays

‡ An appropriate metric for the rotating frame is

$$-ds^2 = dr^2 + r^2\,d\varphi^2 + dz^2 + 2\omega r^2\,d\varphi dt - (c^2 - \omega^2 r^2)dt^2$$

§ A few remarks on the clock paradox are in order. It was noted in Einstein's first paper on the electrodynamics of moving bodies that if we have two identical clocks and keep one at rest in an inertial frame and then move the other in a closed path which returns to the position of the first clock, the two clocks will no longer agree. Consider now a pair of twins. One remains at rest in an inertial frame and the other sets off in a rocket and then returns. The traveler will on returning find that the stay-at-home twin is older than he. Darwin has pointed out (C. G. Darwin, *Nature* **180**, 976 (1957)) that the entire problem may be understood within the framework of special relativity, for the acceleration times of the rocket may be short and the period of uniform motion extremely long. Then the result cannot depend on what happened during the short acceleration periods. If the velocity of the traveler relative to the twin at rest in an inertial frame is v then

$$t_r = t_m / \sqrt{1 - v^2/c^2}$$

where t_r is the elapsed time for the twin at rest and t_m is the elapsed time for the moving twin. This result follows immediately from special relativity, since we must carry out all calculations in the frame of the twin who remains fixed at all times in the inertial frame. The identical result may be obtained (see C. Møller, *The Theory of Relativity*, Oxford University Press, New York, 1952) if we calculate in the frame of the moving twin, using the formalism of general relativity, which is appropriate for frames that may undergo accelerations. There is, therefore, no paradox.

are always geodesics and that a curved-space metric is required in a gravitational field.

It has been pointed out (3, 4) that by suitably redefining, operationally, the measurement of length and time, the Lorentz metric may always be used. We could similarly insist that the earth's surface is flat. By suitably defining the operations of measurement as a function of latitude and longitude, internal consistency would be achieved with Euclidean geometry. This is not the point of view which we adopt in this tract.

2.4 General Covariance

If we consider gravitational fields alone, the equivalence principle denies us the possibility of distinguishing, by local measurements, between an inertial frame and a freely falling system in a gravitational field. There is then no a priori reason to give special significance to inertial frames. Also it is not possible to set up the required system of synchronized clocks throughout a gravitational field. For these reasons Einstein was led to postulate that all systems of coordinates are equally good for the description of nature and that the laws of physics should have the same form in all. This is the principle of general covariance.

If we adopt this principle the coordinates become nothing more than a bookkeeping system to label the events. The principle of general covariance has been a valuable guide in deducing correct equations. It leads us to avoid principles which seem simple only in certain coordinate systems, and to retain those which can be simply expressed in arbitrary systems of coordinates. It has been pointed out by Kretschmann (5) that any physical law can be written in a covariant form. The result is usually not simple. Kretschmann also pointed out that the principle of general covariance therefore has absolutely no necessary physical consequences. The requirements of simplicity of form plus covariance have nonetheless been a valuable guide in deducing equations which in a final analysis must stand or fall on the basis of comparison with experiment.

The treatment of generally covariant equations in a curved space is facilitated by the formalism of the tensor calculus.

References

1. H. Minkowski, address to the 80th Assembly of German Natural Scientists and Physicians, Cologne, 1908; reprinted in *The Principle of Relativity*, Dover Publications, New York, 1952.
2. A. Einstein, *Ann. Physik* **49**, 769 (1916).
3. N. Rosen, *Phys. Rev.* **57**, 147 (1949).
4. R. H. Dicke, *Revs. Modern Phys.* **29**, 363 (1957).
5. E. Kretschmann, *Ann. Physik* **53**, 575 (1917).

CHAPTER 3

Riemannian Geometry and Tensor Calculus

The inner properties of surfaces are "most worthy of being diligently exploited by geometers." C. F. Gauss

3.1 Some Ideas about Curvature

A more general kind of geometry, in which $g_{\mu\nu}$ is not necessarily reducible by a coordinate transformation everywhere to the Lorentz metric, will now be discussed.

Gauss considered the following question. Suppose we have a two-dimensional curved surface, inhabited by intelligent two-dimensional animals. Can they determine that

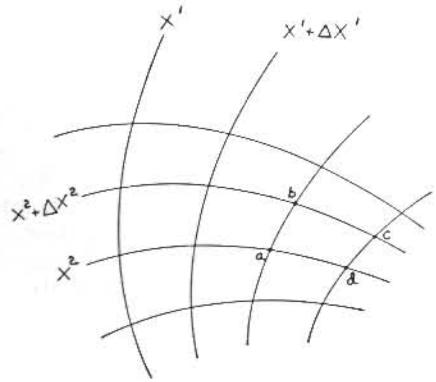

Fig. 3.1

their space is curved? Is it possible to determine the elements of curvature by means of measurements made within the surface alone? He found that this can indeed be done. First we proceed by labeling the points of the surface in any regular but nonetheless arbitrary way. Two arbitrary families of curves, $x^1 = \text{constant}$ and $x^2 = \text{constant}$ are the coordinate system (Fig. 3.1). Direct measurement of length between points a and b gives g_{22}; similarly by measuring lengths ac and ad we obtain g_{12} and g_{11}. Gauss gave formulas which allow the curvature to be written in terms of the $g_{\mu\nu}$, and their derivatives.

Curvature is an intrinsic property, at any given point the same value is obtained in every coordinate system. We shall see how the idea of curvature can be extended to more than two dimensions. Einstein's theory of gravitation relates the curvature of the space to the distribution of stress and energy. This follows in part the suggestion of Mach to the effect that the properties of the space-time continuum are determined by the distribution of energy.

3.2 Transformation Laws for Different Kinds of Tensors

Let us start our discussion with the assumption that we have M variables $x^1, x^2, x^3, \ldots, x^M$. A set of particular values of these variables is now regarded as a point in a hyperspace or manifold having M dimensions. The space is made up of all the points corresponding to the range of values which can be assumed by these variables. Suppose we employ a different labeling for the points $x'^1, x'^2, x'^3, \ldots, x'^M$, such that

$$x'^\alpha = f^\alpha(x^1, x^2, \ldots, x^M) \tag{3.1}$$

We assume that derivatives exist, and write

$$dx'^\alpha = \frac{\partial f^\alpha}{\partial x^\beta} dx^\beta = \frac{\partial x'^\alpha}{\partial x^\beta} dx^\beta \tag{3.2}$$

The coordinate differentials dx^α are said to be the components of a contravariant vector.‡ Similarly, any set of quantities F^α are defined to be a contravariant vector if they obey the transformation law

‡ A vector is a tensor of the first rank, a scalar is a tensor of rank zero.

$$F'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta} F^\beta \tag{3.3}$$

Consider now quantities such as $\partial\varphi/\partial x^\beta$, where φ is some function of the variables $x^1, x^2, \ldots, x^M$:

$$\frac{\partial \varphi}{\partial x'^\alpha} = \frac{\partial \varphi}{\partial x^\beta}\frac{\partial x^\beta}{\partial x'^\alpha} \tag{3.4}$$

The quantities $\partial\varphi/\partial x^\alpha$ are seen to obey a different transformation law from (3.3) and any set of quantities transforming according to

$$K'_\alpha = \frac{\partial x^\beta}{\partial x'^\alpha} K_\beta \tag{3.5}$$

are said to form a covariant‡ vector.

Note that our definitions of covariant and contravariant vectors require the existence of derivatives on our manifold and not the existence of a metric. We are following the convention of denoting covariant vectors by subscripts and contravariant vectors by superscripts.

The product of two contravariant vectors A^α and B^β will satisfy the transformation law

$$A'^\alpha B'^\beta = \frac{\partial x'^\alpha}{\partial x^\nu}\frac{\partial x'^\beta}{\partial x^\delta} A^\nu B^\delta \tag{3.6}$$

A set of quantities $T^{\mu\nu}$, which obey the transformation law (3.6), are said to form a contravariant tensor of the second rank. Similarly a covariant second-rank tensor is one which obeys the transformation law

$$T'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} T_{\alpha\beta} \tag{3.7}$$

A mixed tensor of any rank obeys the transformation

‡ The word covariant has two quite different meanings. A covariant theory or equation has the same form in all systems of coordinates. The word covariant is used also to signify that a tensor obeys the transformation law (3.5). Thus a covariant equation may contain contravariant tensors as well as covariant ones, and other objects which are not tensors.

law

$$T'^{\mu\nu\cdots}_{\epsilon\eta\cdots} = \frac{\partial x^\alpha}{\partial x'^\epsilon}\frac{\partial x^\beta}{\partial x'^\eta}\cdots\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}\cdots T_{\alpha\beta}{}^{\rho\sigma} \qquad (3.8)$$

Certain other quantities transform according to the law

$$T'^{\mu\cdots}_{\epsilon\cdots} = J^W \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\epsilon}\cdots T^{\alpha\cdots}_{\beta\cdots} \qquad (3.9)$$

J is the Jacobian determinant $|\partial x^\kappa/\partial x'^\iota|$. The superscript W is the power to which J is raised. $T^{\mu\cdots}_{\nu\cdots}$ is said to be a tensor density of weight W.

A function S, which transforms to S', such that $S = S'$ at every point and in all coordinate systems is said to be an invariant, or scalar.

In all cases when a quantity is given, its form provides a prescription for getting it in other coordinate systems. Thus if $T_{\mu\nu} = A_\mu B_\nu$, then in a different system of coordinates $T'_{\mu\nu} = A'_\mu B'_\nu$.

The product $A_\lambda B^\lambda$ of a covariant and a contravariant vector transforms in the following way:

$$A'_\lambda B'^\lambda = \frac{\partial x^\mu}{\partial x'^\lambda}\frac{\partial x'^\lambda}{\partial x^\nu} A_\mu B^\nu = \delta_\nu{}^\mu A_\mu B^\nu = A_\nu B^\nu \qquad (3.10)$$

This product is therefore a scalar. It also follows that the inner product (summation over upper-lower index pairs) is a scalar for tensors of higher rank. These notions may be employed to test for tensor character. Let B^μ be an arbitrary contravariant tensor and let A_μ be a set of quantities which may or may not have tensor character. Then if the product $A_\mu B^\mu$ is an invariant we can show that A_μ is a tensor, for

$$A_\mu B^\mu = A'_\mu B^\alpha \frac{\partial x'^\mu}{\partial x^\alpha} \qquad (3.11)$$

An index such as α or μ, over which a sum is to be carried out, may be given any convenient letter. This is an important aid in manipulation. We may then write (3.11) as

$$\left(A_\mu - A'_\nu \frac{\partial x'^\nu}{\partial x^\mu}\right) B^\mu = 0 \qquad (3.12)$$

It follows from (3.12) that A_μ transforms according to $A_\mu = A'_\nu \partial x'^\nu/\partial x^\mu$ and is therefore a covariant vector.

The Kronecker delta, written as $\delta_\mu{}^\nu$, is a quantity which is unity if $\mu = \nu$ and zero if $\mu \neq \nu$. If we write

$$\delta_\mu{}^\nu = \frac{\partial x^\nu}{\partial x^\mu} = \frac{\partial x^\nu}{\partial x'^\alpha}\frac{\partial x'^\alpha}{\partial x^\mu} = \frac{\partial x^\nu}{\partial x'^\beta}\frac{\partial x'^\alpha}{\partial x^\mu}\delta'_\alpha{}^\beta \qquad (3.13)$$

it is evident that $\delta_\mu{}^\nu$ is a mixed tensor.

If a tensor $S^{\mu\nu\alpha} = S^{\nu\mu\alpha}$, it is said to be symmetric in the indices μ and ν. If $A^{\alpha\beta\gamma} = -A^{\beta\alpha\gamma}$, the tensor is said to be antisymmetric (or skew-symmetric) in the indices α and β. By writing out the transformation laws it becomes clear that the symmetry properties of a tensor are retained under coordinate transformations if the pair of indices are both subscripts or both superscripts. In general the symmetry properties are not retained if one index is a subscript and one a superscript. The symmetry properties are therefore meaningful only for the same kind of indices.

We may write any tensor $A^{\mu\nu}{}_\alpha{}^\delta$ as

$$A^{\mu\nu}{}_\alpha{}^\delta = \tfrac{1}{2}[A^{\mu\nu}{}_\alpha{}^\delta + A^{\nu\mu}{}_\alpha{}^\delta] + \tfrac{1}{2}[A^{\mu\nu}{}_\alpha{}^\delta - A^{\nu\mu}{}_\alpha{}^\delta] \qquad (3.14)$$

and it follows from (3.14) that any tensor may be considered as the sum of a part which is symmetric and a part which is antisymmetric in a given pair of upper or lower indices.

It follows from the transformation laws that if all components of a given tensor vanish in one coordinate system, then all components vanish in all coordinate systems. This fact is of great importance in theoretical physics. If a law is written in tensor form, for example by saying that one tensor equals another, the difference of the two tensors will vanish in all coordinate systems and the law has a validity independent of the coordinates which may be employed. Similarly, if we establish a tensor equation in a special coordinate system, it is valid in general.