# Tensors, Differential Forms, and Variational Principles

## David Lovelock and Hanno Rund

The aim of this book is to present a self-contained, reasonably modern account of tensor analysis and the calculus of exterior differential forms, adapted to the needs of physicists, engineers and applied mathematicians. In the later, increasingly sophisticated chapters, the interaction between the concept of invariance and the calculus of variations is examined. This interaction is of profound importance to all physical field theories.

Beginning with simple physical examples, the theory of tensors and forms is developed by a process of successive abstraction. This enables the reader to infer generalized principles from concrete situations—departing from the traditional approach to tensors and forms in terms of purely differential-geometric concepts.

The treatment of the calculus of variations of single and multiple integrals is based *ab initio* on Carathéodory's method of equivalent integrals. Subsequent material explores the effects of invariance postulates on variational principles, focusing ultimately on relativistic field theories. Other discussions include:

* integral invariants
* simple and direct derivations of Noether's theorems
* Riemannian spaces with indefinite metrics

The emphasis in this book is on analytical techniques, with abundant problems, ranging from routine manipulative exercises to technically difficult problems encountered by those using tensor techniques in research activities. A special effort has been made to collect many useful results of a technical nature, not generally discussed in the standard literature. The Appendix, newly revised and enlarged for the Dover edition, presents a reformulation of the principal concepts of the main text within the terminology of current global differential geometry, thus bridging the gap between classical tensor analysis and the fundamentals of more recent global theories.

See every Dover book
in print at
**www.doverpublications.com**

ISBN 0-486-65840-6

90000

$15.95 IN USA
$23.95 IN CANADA

9 780486 658407
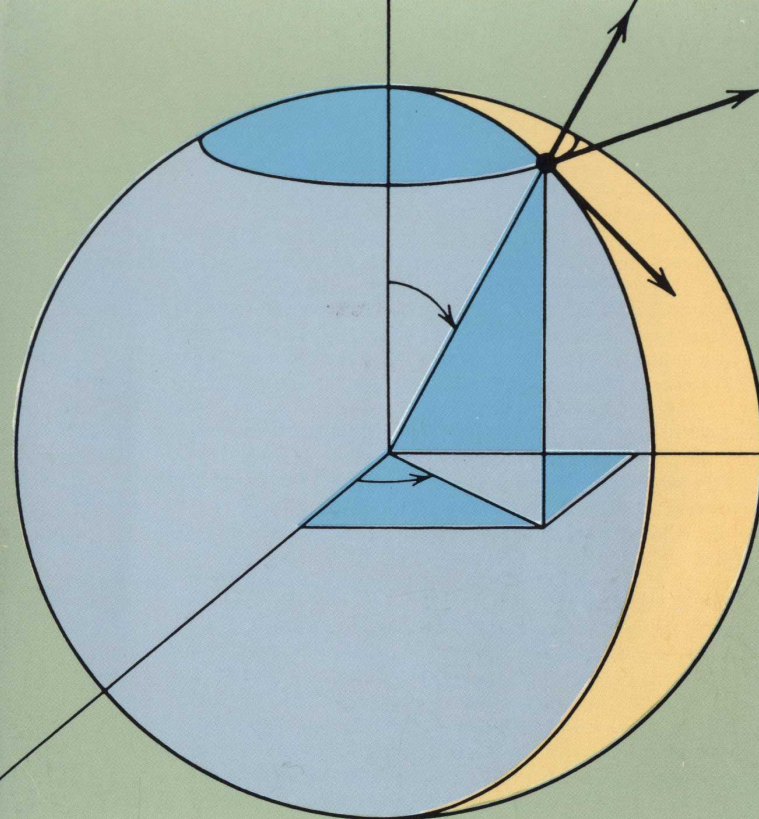
## Tensors, Differential Forms, and Variational Principles

### David Lovelock and Hanno Rund

# TENSORS, DIFFERENTIAL FORMS, AND VARIATIONAL PRINCIPLES

by David Lovelock

*Professor of Mathematics*
*University of Arizona*

and Hanno Rund

*Professor of Mathematics*
*University of Arizona*
*and Adjunct Professor of Applied Mathematics*
*University of Waterloo (Ontario)*

# PREFACE

The objective of this book is twofold. Firstly, it is our aim to present a self-contained, reasonably modern account of tensor analysis and the calculus of exterior differential forms, adapted to the needs of physicists, engineers, and applied mathematicians in general. Secondly, however, it is anticipated that a substantial part of the material included in the later chapters is of interest also to those who have some previous knowledge of tensors and differential forms: we refer in particular to the remarkable interaction between the concept of invariance and the calculus of variations, which has profound implications in almost all physical field theories.

The general approach and organization of the opening chapters is determined almost exclusively by the requirements of our first objective. In fact, these chapters are based on courses presented during the past decade at several universities to audiences with widely varying interests and academic backgrounds. Accordingly the prerequisites consist merely of basic linear algebra, advanced calculus of several real variables, and some very elementary aspects of the theory of differential equations; the first five chapters therefore constitute a one-semester course which should be readily accessible to senior undergraduates majoring in mathematics, physics, or some branch of engineering. Initially the pace is very gradual, if not leisurely, with emphasis on motivation with the aid of simple physical examples, while at the same time a systematic effort is made to proceed to the core of the subject matter by way of successive abstraction from concrete situations. In this respect, therefore, our treatment does not follow the historical development of the calculus of tensors and forms, whose origins are deeply rooted in differential geometry, which does, in fact, provide their most natural setting. However, since the applications of tensors and forms have meanwhile spread to entirely different areas, the basic approach presented below has intentionally been divested of many of the customary trappings of metric differential geometry.

The later parts of the book, beginning with Chapter 6, are devoted primarily to the second of the aforementioned objectives. It had been observed repeatedly by David Hilbert that the effects of invariance postulates on

variational principles are of a surprisingly profound and far-reaching nature, particularly insofar as relativistic field theories are concerned. Since no previous knowledge of the calculus of variations is presupposed, this subject is developed ab initio for single as well as multiple integrals, with special emphasis on the various invariance requirements which may be imposed on the fundamental (or action) integral. Much of this treatment is based on the relatively recent direct methods of Constantin Carathéodory, instead of on the classical theory of the first and second variations of the fundamental integral; thus the approach of Carathéodory, which is undoubtedly more illuminating and powerful than the traditional procedure, should be rendered accessible to a wide class of readers.†

A very important instance of the consequences of the imposition of special invariance requirements on variational principles is represented by the famous theorems of Noether, which, when applied to physical field theories, predict the existence and precise nature of conservation laws. Instead of following the original procedure of Emmy Noether, which is based on some deep and conceptually difficult results in the calculus of variations, it will be seen that these theorems can be derived directly and almost effortlessly by means of elementary tensorial techniques.

In view of the fact that many applications of these ideas are concerned with the general theory of relativity, a chapter on Riemannian spaces with indefinite metrics is included; it should be emphasized, however, that it is only at this stage that the concept of a metric is introduced and used in a systematic manner. Thus, on the one hand, the aesthetic appeal of metric differential geometry, absent from the earlier chapters, is recaptured, while on the other hand, the prerequisites for the final chapter are established. This chapter is considerably more specialized than the remainder of the book: it contains, inter alia, fairly recent results which are of interest primarily to relativists.

To a great extent the emphasis in this book is on analytical techniques. Thus a large number of problems is included, ranging from routine manipulative exercises to technically difficult problems of the kind frequently encountered by those who use tensor techniques in the course of their research activities. Indeed, some considerable trouble has been taken to collect many useful results of a purely technical nature, which generally are not discussed in the standard literature, but which form part of the almost indispensable "folklore" known to most experts in the field. Despite this emphasis on technique, however, every effort has been made to maintain an acceptable level of rigor commensurate with the classical background on

† Incidentally, it should be pointed out that this approach also plays an increasingly significant role in the theory of optimal control (which is not touched upon in this book).

which the analysis is based. Moreover, the modern, more sophisticated and abstract approach to the theory of tensors and forms on manifolds, which is less dependent on the use of coordinate systems, is discussed briefly in the Appendix, which, it is hoped, will assist the reader in bridging the deep chasm between classical tensor analysis and the fundamentals of more recent global theories.

DAVID LOVELOCK
HANNO RUND

*Tucson, Arizona*
*January 1975*

# SUGGESTIONS FOR THE GENERAL USE OF THE BOOK

Chapters 1–5 constitute a one-semester senior undergraduate or junior graduate course. Sections 4.3, 4.4, 5.4, and 5.7 may conceivably be omitted since the remainder of this part of the book is independent of their contents.

Chapters 6–8 constitute a one-semester graduate course, in which Sections 6.4, 6.6, 7.4, 7.5, and 7.6 may be omitted.

An introductory course, specifically designed for the needs of relativists, could be based on Chapters 1–3, together with Sections 4.1, 4.2, 5.1, 5.2, 5.3, 5.4, 5.6, 6.1, 6.2, 7.1, 7.2, and 7.3.

References to equations are of the form (N.M.P), where N and M indicate the corresponding chapter and section, respectively. If N coincides with the chapter at hand it is omitted.

Difficult problems are marked with an asterisk; in some cases explicit bibliographical references are given. No attempt was made to compile an exhaustive bibliography; the latter consists entirely of general references and individual papers cited in the text.

# CONTENTS

# TENSORS,
# DIFFERENTIAL FORMS,
# AND VARIATIONAL
# PRINCIPLES

# 1

# PRELIMINARY

# OBSERVATIONS

One of the principal advantages of classical vector analysis derives from the fact that it enables one to express geometrical or physical relationships in a concise manner which does not depend on the introduction of a coordinate system. However, for many purposes of pure and applied mathematics the concept of a vector is too limited in scope, and to a very significant extent, the tensor calculus provides the appropriate generalization. It, too, possesses the advantage of a concise notation, and the formulation of its basic definitions is such as to allow for effortless transitions from a given coordinate system to another, while in general the inspection of any relation involving tensors permits an inference to be drawn immediately as to whether or not that relation is valid in all allowable coordinate systems.

The objective of this chapter is essentially motivational. A few simple physical and geometrical situations are briefly described in order to reveal the inadequacy of the vector concept under certain circumstances. In particular, whereas in a three-dimensional space a vector is uniquely determined by its three components relative to some coordinate system, there exist important physical and geometrical entities which require more than three components for their complete specification. Tensors are examples of such quantities; however, the definition of the tensor concept is deferred until Chapters 2 and 3, and accordingly any reader who is reasonably familiar with the elementary ideas touched upon below may confine himself to a very casual survey of the contents of this chapter.

## 1.1 SIMPLE EXAMPLES OF TENSORS IN PHYSICS AND GEOMETRY

Although it is not feasible to define the concept of a tensor at this stage, we briefly describe a few elementary examples which serve to indicate quite

clearly that, in the case of many basic geometrical or physical applications, it is necessary to introduce quantities which are more general than vectors. Whereas a vector possesses three components in a three-dimensional space, we shall be confronted with entities which possess more than three components in such spaces.

The following notation is adopted in this section. Suppose that, in a Euclidean space $E_3$, we are given a rectangular coordinate system with origin at a fixed point $O$ of $E_3$. The coordinates of an arbitrary point $P$ of $E_3$ relative to this coordinate system are denoted by $(x^1, x^2, x^3)$, and the unit vectors in the directions of the positive $Ox^1$-, $Ox^2$-, and $Ox^3$-axes are represented by $e_1$, $e_2$, and $e_3$. The latter form a basis of $E_3$ in the sense that any vector $A$ can be expressed in the form

$$A = A_1 e_1 + A_2 e_2 + A_3 e_3, \tag{1.1}$$

in which $A_1, A_2, A_3$ denote the three components of $A$, these being the lengths of the projections of $A$ onto $e_1, e_2, e_3$, respectively. The length $|A|$ of $A$ is given by

$$|A|^2 = (A_1)^2 + (A_2)^2 + (A_3)^2. \tag{1.2}$$

In particular, for the position vector $r$ of $P$ we have

$$r = x^1 e_1 + x^2 e_2 + x^3 e_3, \tag{1.3}$$

and, writing $r = |r|$ for the sake of brevity,

$$r^2 = (x^1)^2 + (x^2)^2 + (x^3)^2. \tag{1.4}$$

EXAMPLE 1    THE STRESS TENSOR OF ELASTICITY

Let us consider an arbitrary elastic medium which is at rest subject to the action of body and/or external loading forces, so that in its interior there will exist internal forces. The stress vector across a surface element $\Delta S$ in the interior of the body at a point $P$ is defined by the limit $\Delta F/\Delta S$ as $\Delta S \to 0$, where $\Delta F$ denotes the resultant of the internal forces distributed over the surface element $\Delta S$. More precisely, let us consider a small cube of the elastic material, whose edges are parallel to the coordinate axes. The faces of the cube whose outward normals are $e_1$, $e_2$, $e_3$, are respectively denoted by $S_1, S_2, S_3$. Let $F_{(k)}$ denote the stress over $S_{(k)}$ ($k = 1, 2, 3$) (see Figure 1). Each of these vectors can be decomposed into components in accordance with (1.1):

$$F_{(k)} = \tau_{k1} e_1 + \tau_{k2} e_2 + \tau_{k3} e_3, \tag{1.5}$$

where $\tau_{kj}$ represents the $j$th component ($j = 1, 2, 3$) of the stress over the face $S_k$. The quantities $\tau_{kj}$ are called the *components of the stress tensor*: they completely specify the internal forces of the elastic medium at each point.
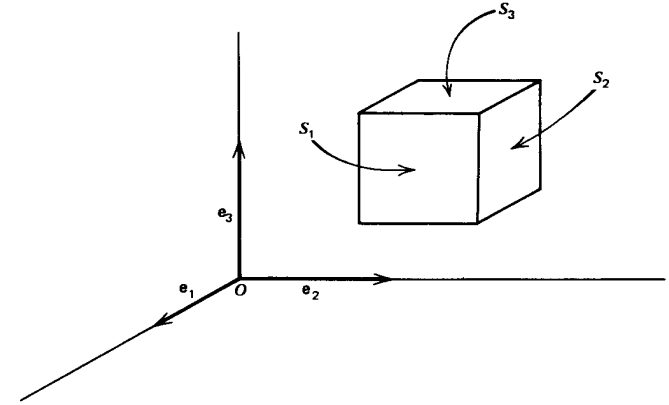
**Fig. 1**

It should be emphasized that this description depends on a system of nine components $\tau_{kj}$. (These are not generally independent but satisfy the symmetry condition $\tau_{kj} = \tau_{jk}$ provided that certain conditions are satisfied, but this phenomenon is not relevant to this discussion.)

As a result of the forces acting on the body, a deformation of the medium occurs, which is described as follows. A point $P$ of the medium, whose initial coordinates are $(x^1, x^2, x^3)$, will be transferred to a position with coordinates $(y^1, y^2, y^3)$, and we shall write $u^j = y^j - x^j$ ($j = 1, 2, 3$). If it is assumed that $u^j$ is a continuously differentiable function of position, one may define the quantities

$$s_{jk} = \frac{1}{2}\left(\frac{\partial u^j}{\partial x^k} + \frac{\partial u^k}{\partial x^j}\right), \tag{1.6}$$

which represent the components of the *strain tensor*. (We do not motivate this definition here, the reader being referred to texts on the theory of elasticity, e.g., Green and Zerna [1].) Again we remark that the specification of $s_{jk}$ depends on nine components (apart from symmetry). Moreover, one of the basic physical assumptions of the theory is contained in the so-called generalized *Hooke's law*, which states that there exists a linear relationship between the components of stress and the components of strain. This is expressed in the form

$$\tau_{jk} = \sum_{l=1}^{3} \sum_{m=1}^{3} c_{jklm} s_{lm} \qquad (j, k = 1, 2, 3), \tag{1.7}$$

in which the coefficients $c_{jklm}$ are of necessity four-index symbols (as contrasted with ordinary vectors, whose components are merely one-index symbols).

It should be abundantly clear from this very cursory description that ordinary vector theory is totally inadequate for the purposes of an analysis of this kind.

### EXAMPLE 2    THE INERTIA TENSOR OF A RIGID BODY

Let us consider a rigid body rotating about a fixed point $O$ in its interior. The angular momentum about $O$ of a particle of mass $m$ of the body located at a point $P$ with position vector $\mathbf{r}$ relative to $O$ is given by $\mathbf{r} \times \mathbf{p}$, where $\mathbf{p} = m\, d\mathbf{r}/dt$ is the linear momentum of the particle. If $\boldsymbol{\omega}$ denotes the angular velocity vector we have that $d\mathbf{r}/dt = \boldsymbol{\omega} \times \mathbf{r}$, and consequently the angular momentum of the particle about $O$ is $m[\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})]$. Thus if the mass density of the body is $\rho$, its total angular momentum about $O$ is given by the integral

$$\mathbf{H} = \int \rho [\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r})]\, dV, \tag{1.8}$$

in which $dV$ denotes the volume element, and where the integration is to be performed over the entire rigid body, (e.g., see Goldstein [1]).

In order to obtain a more useful expression for $\mathbf{H}$, we introduce a rectangular coordinate system with basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ and origin at $O$, this system being fixed relative to the rigid body. In terms of (1.3) we have, by definition of vector product,

$$\boldsymbol{\omega} \times \mathbf{r} = \mathbf{e}_1(\omega_2 x^3 - \omega_3 x^2) + \mathbf{e}_2(\omega_3 x^1 - \omega_1 x^3) + \mathbf{e}_3(\omega_1 x^2 - \omega_2 x^1),$$

and, after a little simplification,

$$\begin{aligned}
\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) = {} & \mathbf{e}_1\{\omega_1[(x^2)^2 + (x^3)^2] - \omega_2 x^1 x^2 - \omega_3 x^1 x^3\} \\
& + \mathbf{e}_2\{-\omega_1 x^1 x^2 + \omega_2[(x^3)^2 + (x^1)^2] - \omega_3 x^2 x^3\} \\
& + \mathbf{e}_3\{-\omega_1 x^1 x^3 - \omega_2 x^2 x^3 + \omega_3[(x^1)^2 + (x^2)^2]\}. \tag{1.9}
\end{aligned}$$

This expression is to be substituted in (1.8). Since the angular velocity components $\omega_1, \omega_2, \omega_3$ are independent of position, the following integrals will occur, for which a special notation is introduced:

$$\left. \begin{aligned}
I_{11} &= \int \rho[(x^2)^2 + (x^3)^2]\, dV, \qquad I_{22} = \int \rho[(x^3)^2 + (x^1)^2]\, dV, \\
I_{33} &= \int \rho[(x^1)^2 + (x^2)^2]\, dV,
\end{aligned} \right\} \tag{1.10}$$

together with

$$\left. \begin{aligned}
I_{12} &= -\int \rho x^1 x^2\, dV = I_{21}, \qquad I_{13} = -\int \rho x^1 x^3\, dV = I_{31}, \\
I_{23} &= -\int \rho x^2 x^3\, dV = I_{32}.
\end{aligned} \right\} \tag{1.11}$$

In terms of this notation, then, the substitution of (1.9) in (1.8) yields the following expression for the total angular momentum of our rigid body about the point $O$:

$$\begin{aligned}
\mathbf{H} = {} & \mathbf{e}_1(I_{11}\omega_1 + I_{12}\omega_2 + I_{13}\omega_3) + \mathbf{e}_2(I_{21}\omega_1 + I_{22}\omega_2 + I_{23}\omega_3) \\
& + \mathbf{e}_3(I_{31}\omega_1 + I_{32}\omega_2 + I_{33}\omega_3) \\
= {} & \mathbf{e}_1 \sum_{h=1}^{3} I_{1h}\omega_h + \mathbf{e}_2 \sum_{h=1}^{3} I_{2h}\omega_h + \mathbf{e}_3 \sum_{h=1}^{3} I_{3h}\omega_h = \sum_{j=1}^{3} \sum_{h=1}^{3} I_{jh}\omega_h \mathbf{e}_j. \tag{1.12}
\end{aligned}$$

This is the formula that we have been seeking; it clearly permits the computation of the angular momentum $\mathbf{H}$ for any angular velocity $\boldsymbol{\omega}$ whenever the quantities (1.10), the so-called moments of inertia, together with the quantities (1.11), the so-called coefficients of inertia, are known. However, these quantities are defined uniquely by the structure of the rigid body, the position of the point $O$, and the directions $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ of the axes through $O$. This phenomenon clearly exhibits the fundamentally important role played by the coefficients $I_{jh}$, which are known collectively as the *components of the inertia tensor*. Again it should be stressed that these are two-index quantities.

The inertia tensor may also be used to calculate the total kinetic energy $T$ of the rigid body due to its rotational motion. The kinetic energy of the particle of mass $m$ located at $P$ is $\frac{1}{2}m|d\mathbf{r}/dt|^2$, and by a process of integration similar to that sketched above it is easily shown that

$$T = \frac{1}{2} \sum_{j=1}^{3} \sum_{h=1}^{3} I_{jh}\omega_j \omega_h. \tag{1.13}$$

Let us suppose, for the moment, that the inertia tensor of a rigid body is known for a given set of axes through a point $O$. For a different physical problem involving the same rigid body it may be necessary to determine the inertia tensor for another set of axes through a different center of rotation. Thus it would be desirable to have some knowledge of the behavior of the components of the inertia tensor under a change of rectangular axes. This state of affairs exemplifies the central theme of the tensor calculus, which is, in fact, deeply concerned with the transformation properties of entities of this kind.

EXAMPLE 3    THE VECTOR PRODUCT

Let $\mathbf{A}$, $\mathbf{B}$ be two vectors with a common point of application $P$ in $E_3$, the angle between these vectors (defined by a rotation from $\mathbf{A}$ to $\mathbf{B}$) being denoted by $\theta$. If $\theta \neq 0$ and $\theta \neq \pi$, these vectors span a plane $\Pi$ through $P$. The vector product $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ at $P$ is defined geometrically as follows: (1) the magnitude $|\mathbf{C}|$ of $\mathbf{C}$ is given by $|\mathbf{A}| |\mathbf{B}| \sin \theta$; (2) the vector $\mathbf{C}$ is normal to $\Pi$ in such a manner that $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ form a right-handed system. This definition is equivalent to the analytical expression

$$\mathbf{C} = \mathbf{e}_1(A_2 B_3 - A_3 B_2) + \mathbf{e}_2(A_3 B_1 - A_1 B_3) + \mathbf{e}_3(A_1 B_2 - A_2 B_1). \tag{1.14}$$

It should be remarked that the above geometrical definition of $\mathbf{A} \times \mathbf{B}$ is meaningful solely for the case of a three-dimensional space, because in a higher-dimensional space the plane $\Pi$ does not possess a unique normal. However, the formula (1.14) suggests quite clearly that in an $n$-dimensional space with $n \geq 3$ we should define the quantities

$$C_{jh} = A_j B_h - A_h B_j \qquad (j, h = 1, \ldots, n). \tag{1.15}$$

These two-index symbols represent the components of an entity which is an obvious generalization to $n$ dimensions of the three-dimensional vector product. However, this entity is not a vector unless $n = 3$. The geometrical reason for this is evident from the remark made above; the analytical reason emerges from the fact that the components $C_1$, $C_2$, $C_3$ of $\mathbf{C}$ as given by (1.14) are related to the quantities (1.15) for $n = 3$ by

$$C_1 = C_{23}, \qquad C_2 = C_{31}, \qquad C_3 = C_{12}, \tag{1.16}$$

or

$$C_j = C_{hk}, \tag{1.17}$$

where $j, h, k$ is an even permutation of the integers 1, 2, 3, and this construction is not feasible if the three indices $j, h, k$ are allowed to assume integer values greater than 3. Thus the generalization (1.15) of the vector product is represented by quantities which are not vector components; in fact, it will be seen that (1.15) exemplifies a certain type of tensor.

## 1.2    VECTOR COMPONENTS IN CURVILINEAR COORDINATE SYSTEMS

Let us regard a given vector $\mathbf{A}$, with fixed point of application $P$ in $E_3$, as a directed segment in the usual sense. Relative to a rectangular coordinate system with origin at $O$ and basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, we can represent $\mathbf{A}$ according to (1.1), the length $|\mathbf{A}|$ of $\mathbf{A}$ being given by (1.2). Again, it should be emphasized that the components $A_1$, $A_2$, $A_3$ of $\mathbf{A}$ which appear in (1.1)

are obtained by a process of projection onto the axes, which implies that these components, and hence the expression for the length $|\mathbf{A}|$, are independent of the location of the point of application $P$.

However, when we attempt to carry out a similar program relative to a curvilinear system of coordinates, the position is not quite so simple. Since in this case there does not exist a universal triple of coordinate axes, we cannot define the components of $\mathbf{A}$ by means of a direct process of projection; moreover, it is not immediately evident what form the counterpart of the formula (1.1) should assume. In this connection it should be remarked that (1.1) is not a vector equation of the type $\mathbf{A} = \mathbf{B}$, which is independent of the choice of the coordinate system: on the contrary, it is a nonvectorial relation which depends crucially on the fact that it refers to a rectangular system.

In order to illustrate these remarks we shall consider in the first instance a special type of curvilinear coordinates, namely, a spherical polar coordinate system whose pole is located at the origin $O$ of the given rectangular system and whose polar axis coincides with $\mathbf{e}_3$. The spherical polar coordinates $(\rho, \theta, \phi)$ are related to the rectangular coordinates $(x^1, x^2, x^3)$ according to the transformation

$$\left. \begin{array}{l} x^1 = \rho \sin \theta \cos \phi, \\ x^2 = \rho \sin \theta \sin \phi, \\ x^3 = \rho \cos \theta, \end{array} \right\} \tag{2.1}$$

as is immediately evident from Figure 2. It should be noted that $\rho \geq 0$, $0 \leq \theta \leq \pi, 0 \leq \phi < 2\pi$.

In view of the fact that such a coordinate system does not possess a set of rectangular axes to which any point of $E_3$ can be referred, we now endeavor to construct a special system of axes *at each point* $P$ of $E_3$. To this end we recall that, in the case of a rectangular system, three lines respectively parallel to the $x^1$-, $x^2$-, and $x^3$-axes can be constructed at any point $P$ by the intersections of appropriate pairs of planes through $P$; for instance, the intersection of the planes $x^2 = $ const., $x^3 = $ const. defines a line parallel to the $x^1$-axis. By way of analogy we shall now consider the surfaces $\rho = c_1$, $\theta = c_2$, $\phi = c_3$, which pass through a point $P$ with spherical polar coordinates $c_1, c_2, c_3$ (it being supposed that $c_1 > 0, 0 < c_2 < \pi, 0 < c_3 < 2\pi$). The following should be noted:

1.    The surface $\rho = c_1$ is a sphere of radius $c_1$ centered at $O$.
2.    The surface $\theta = c_2$ is a circular cone with vertex $O$, whose axis coincides with $Ox^3$, and whose semi-angle is $c_2$.
3.    The surface $\phi = c_3$ is a plane containing $Ox^3$, whose intersection with the $(x^1 x^2)$-plane is inclined at an angle $c_3$ to the $Ox^1$ axis.
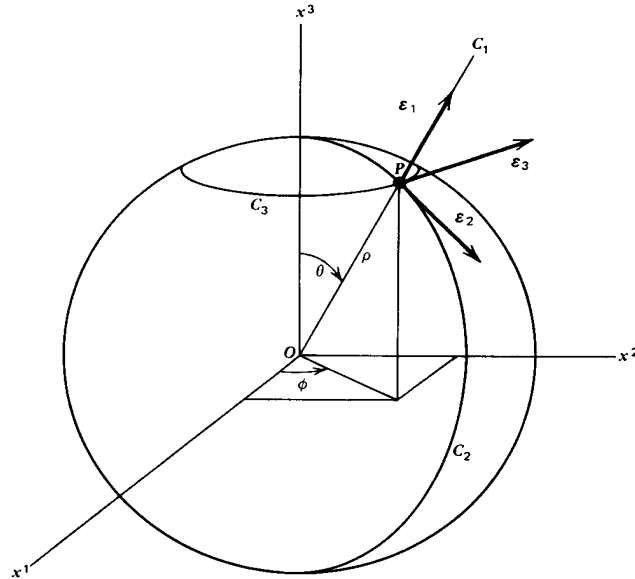
Fig. 2

It then follows:

1.  The $\rho$-coordinate curve, defined to be the intersection of the surfaces 2 and 3, is a straight line issuing from $O$, denoted by $C_1$ in Figure 2.
2.  The $\theta$-coordinate curve, defined to be the intersection of the surfaces 1 and 3, is a great circle on the sphere $\rho = c_1$, denoted by $C_2$ in Figure 2.
3.  The $\phi$-coordinate curve, defined to be the intersection of the surfaces 1 and 2, is the circle on the sphere $\rho = c_1$ containing all points satisfying $\theta = c_2$, this curve being denoted by $C_3$ in Figure 2.

Let us now construct the three unit vectors $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ which are respectively tangent to the coordinate curves $\rho = c_1, \theta = c_2, \phi = c_3$ at $P$, the directions of these vectors being chosen in the directions of increasing $\rho, \theta, \phi$. From Figure 2 it is easily verified that relative to our rectangular coordinates the components of $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ are given by

$$\left.\begin{array}{l} \boldsymbol{\varepsilon}_1 = (\sin\theta\cos\phi, \ \sin\theta\sin\phi, \ \cos\theta), \\ \boldsymbol{\varepsilon}_2 = (\cos\theta\cos\phi, \ \cos\theta\sin\phi, \ -\sin\theta), \\ \boldsymbol{\varepsilon}_3 = (-\sin\phi, \ \cos\phi, \ 0). \end{array}\right\} \qquad (2.2)$$

Our construction has thus given rise to a set of three linearly independent vectors at the point $P$. It is immediately evident that this construction can be carried out at any point, and hence *our spherical polar coordinate system defines a basis at each point of $E_3$*. However, *the directions of the corresponding basis vectors* $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ *vary from point to point*, for, according to (2.2), their components relative to our fixed rectangular coordinate system depend on the polar coordinates $\theta, \phi$ of the point under consideration. It will be seen presently that this idea of a "moving frame of reference" is of paramount importance to all subsequent developments.

Now let us consider the vector $\mathbf{A}$ whose point of application is located at $P$. At first sight one might be inclined to define the components of $\mathbf{A}$ relative to the spherical polar coordinate system to be the oriented lengths of the projections of $\mathbf{A}$ onto the basis vectors $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ at $P$. However, a little reflection shows that a somewhat more sophisticated approach is required. Let $d\mathbf{r}$ denote an arbitrary increment of the position vector $\mathbf{r}$ of $P$. Relative to the rectangular coordinate system the components of $d\mathbf{r}$ are simply $dx^1, dx^2, dx^3$, which is in accordance with (1.3), and thus it is natural to *define* the components of $d\mathbf{r}$ relative to the spherical polar coordinate system to be $d\rho, d\theta, d\phi$, where $\rho + d\rho, \theta + d\theta, \phi + d\phi$ denote the coordinates of the point whose position vector is $\mathbf{r} + d\mathbf{r}$. However, it is evident from Figure 3 that the oriented lengths of the projections of $d\mathbf{r}$ onto $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ are respectively $d\rho, \rho\, d\theta, \rho\sin\theta\, d\phi$, so that

$$d\mathbf{r} = d\rho\,\boldsymbol{\varepsilon}_1 + \rho\, d\theta\,\boldsymbol{\varepsilon}_2 + \rho\sin\theta\, d\phi\,\boldsymbol{\varepsilon}_3. \qquad (2.3)$$

Thus the coefficients of $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3$ are not merely the components $(d\rho, d\theta, d\phi)$ of $d\mathbf{r}$, but rather these components multiplied respectively by the so-called *scale factors* $(1, \rho, \rho\sin\theta)$, which are functions of the positional coordinates of $P$.

Returning to the arbitrary vector $\mathbf{A}$ at $P$, we note that it is always possible to construct a curve $\mathbf{r} = \mathbf{r}(t)$ through $P$, which is such that, for a suitable choice of its parameter $t$, the tangent vector $d\mathbf{r}/dt$ is identical with $\mathbf{A}$ at $P$. In accordance with the above definition, the components of $d\mathbf{r}/dt$ relative to the spherical polar coordinate system are $(d\rho/dt, d\theta/dt, d\phi/dt)$; obviously these quantities are also the corresponding components $\overline{A}_1, \overline{A}_2, \overline{A}_3$ of $\mathbf{A}$ at $P$, the notation $\overline{A}_j$ having been chosen for the purpose of distinguishing these components from the components $A_j$ of $\mathbf{A}$ relative to our rectangular coordinate system. By virtue of our construction we have from (2.3):

$$\mathbf{A} = \frac{d\mathbf{r}}{dt} = \frac{d\rho}{dt}\,\boldsymbol{\varepsilon}_1 + \rho\frac{d\theta}{dt}\,\boldsymbol{\varepsilon}_2 + \rho\sin\theta\frac{d\phi}{dt}\,\boldsymbol{\varepsilon}_3,$$

or, in terms of the above identification,
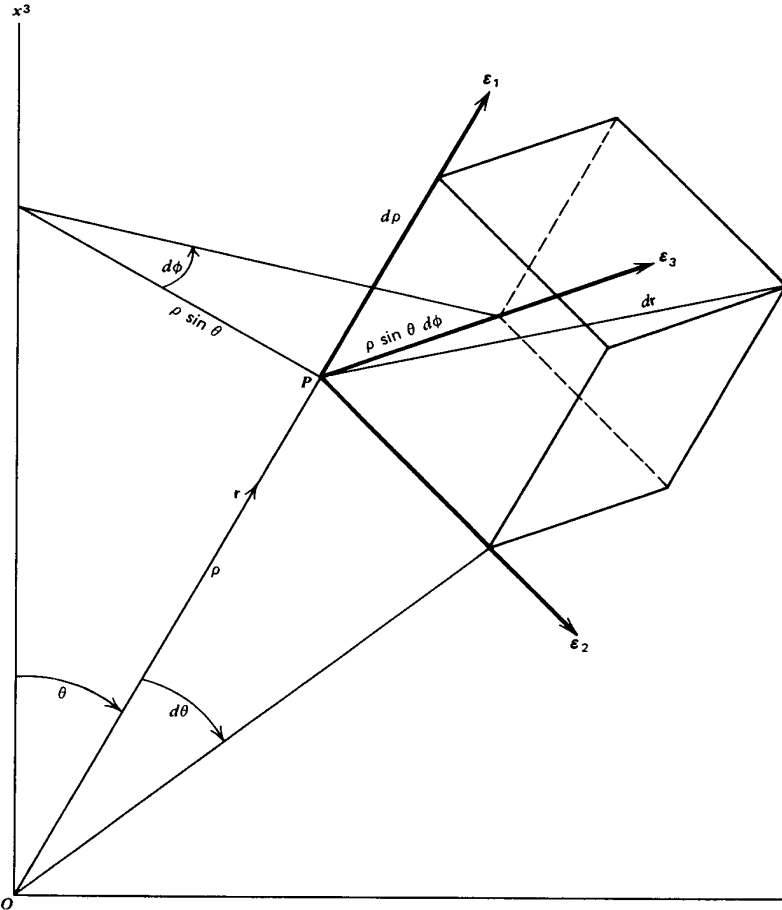
**Fig. 3**

$$\mathbf{A} = \bar{A}_1 \boldsymbol{\varepsilon}_1 + \rho \bar{A}_2 \boldsymbol{\varepsilon}_2 + \rho \sin \theta \, \bar{A}_3 \boldsymbol{\varepsilon}_3, \tag{2.4}$$

in which the scale factors $(1, \rho, \rho \sin \theta)$ appear once more.

It is easily verified that the basis vectors of our spherical polar coordinate system are mutually orthogonal:

$$\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2 = 0, \qquad \boldsymbol{\varepsilon}_2 \cdot \boldsymbol{\varepsilon}_3 = 0, \qquad \boldsymbol{\varepsilon}_3 \cdot \boldsymbol{\varepsilon}_1 = 0. \tag{2.5}$$

(It should be emphasized, however, that this property is not common to all curvilinear coordinate systems.) From (2.4) it therefore follows that

$$\mathbf{A} \cdot \boldsymbol{\varepsilon}_1 = \bar{A}_1, \qquad \mathbf{A} \cdot \boldsymbol{\varepsilon}_2 = \rho \bar{A}_2, \qquad \mathbf{A} \cdot \boldsymbol{\varepsilon}_3 = \rho \sin \theta \bar{A}_3, \tag{2.6}$$

from which $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ may be evaluated directly with the aid of (2.2). It is found that

$$\left. \begin{aligned}
\bar{A}_1 &= \sin \theta \cos \phi A_1 + \sin \theta \sin \phi A_2 + \cos \theta A_3, \\
\bar{A}_2 &= \frac{1}{\rho} \cos \theta \cos \phi A_1 + \frac{1}{\rho} \cos \theta \sin \phi A_2 - \frac{1}{\rho} \sin \theta A_3, \\
\bar{A}_3 &= -\frac{\sin \phi}{\rho \sin \theta} A_1 + \frac{\cos \phi}{\rho \sin \theta} A_2.
\end{aligned} \right\} \tag{2.7}$$

We remark that this represents a *linear relationship between the components of* $\mathbf{A}$ *in the respective coordinate systems*, which can be expressed in the form

$$\bar{A}_j = \sum_{h=1}^{3} \alpha_{jh} A_h, \tag{2.8}$$

in which, however, *the coefficients* $\alpha_{jh}$ *are functions of the positional coordinates* $(\rho, \theta, \phi)$ *of the point P of application of* $\mathbf{A}$.

In passing we note that we can express the length $|\mathbf{A}|$ of the vector $\mathbf{A}$ at $P$ in terms of its components (2.7). For, from (2.4) and (2.5) it follows directly that

$$|\mathbf{A}|^2 = (\mathbf{A} \cdot \mathbf{A}) = (\bar{A}_1)^2 + \rho^2 (\bar{A}_2)^2 + \rho^2 \sin^2 \theta (\bar{A}_3)^2. \tag{2.9}$$

Thus when spherical polar coordinates are used, the square of the length of a vector is not given by the sum of the squares of its components, but by a quadratic form in these components *whose coefficients depend on the position of the point of application of the vector.*

We shall now briefly consider an arbitrary curvilinear coordinate system in which the coordinates of a point $P$ are denoted by $(u, v, w)$. Our considerations will be restricted to a region of $E_3$ in which $P$ is uniquely represented by this triple, so that there exist three relations which specify the rectangular coordinates $x^1, x^2, x^3$ of $P$ uniquely as functions of $u, v, w$:

$$x^1 = f^1(u, v, w), \qquad x^2 = f^2(u, v, w), \qquad x^3 = f^3(u, v, w). \tag{2.10}$$

Again, if **r** denotes the position vector of $P$, we can represent (2.10) vectorially as

$$\mathbf{r} = \mathbf{r}(u, v, w). \tag{2.11}$$

The following construction is a direct generalization of the procedure sketched above. The $u$-coordinate curves are defined by the intersections of the surfaces $v = $ const., $w = $ const., etc., and the tangent vectors to the $u$-, $v$-, $w$-coordinate curves are respectively given by $\mathbf{r}_u = d\mathbf{r}/du$, $\mathbf{r}_v = d\mathbf{r}/dv$, $\mathbf{r}_w = d\mathbf{r}/dw$ (by definition of partial derivative). The unit vectors in the directions of these tangents are again denoted by $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\boldsymbol{\varepsilon}_3$, so that

$$\boldsymbol{\varepsilon}_1 = \frac{\mathbf{r}_u}{|\mathbf{r}_u|}, \qquad \boldsymbol{\varepsilon}_2 = \frac{\mathbf{r}_v}{|\mathbf{r}_v|}, \qquad \boldsymbol{\varepsilon}_3 = \frac{\mathbf{r}_w}{|\mathbf{r}_w|}, \tag{2.12}$$

where the normalization factors $|\mathbf{r}_u|^{-1}$, $|\mathbf{r}_v|^{-1}$, $|\mathbf{r}_w|^{-1}$ are to be evaluated directly from (2.11). Thus at each point of the region of $E_3$ under consideration a basis is defined, it being assumed that the equations (2.10) are such that the vectors $\mathbf{r}_u$, $\mathbf{r}_v$, $\mathbf{r}_w$ are linearly independent. Also, from (2.11) we have

$$d\mathbf{r} = du\,\mathbf{r}_u + dv\,\mathbf{r}_v + dw\,\mathbf{r}_w, \tag{2.13}$$

in which $du$, $dv$, $dw$ are the components of $d\mathbf{r}$ relative to our curvilinear coordinate system. Thus if we substitute from (2.12), we find that

$$d\mathbf{r} = |\mathbf{r}_u|\,du\,\boldsymbol{\varepsilon}_1 + |\mathbf{r}_v|\,dv\,\boldsymbol{\varepsilon}_2 + |\mathbf{r}_w|\,dw\,\boldsymbol{\varepsilon}_3, \tag{2.14}$$

so that the scale factors are simply $|\mathbf{r}_u|$, $|\mathbf{r}_v|$, $|\mathbf{r}_w|$. In accordance with our previous construction we now define the components $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ of the vector **A** at $P$ relative to our curvilinear coordinate system by the relation

$$\mathbf{A} = |\mathbf{r}_u|\bar{A}_1\boldsymbol{\varepsilon}_1 + |\mathbf{r}_v|\bar{A}_2\boldsymbol{\varepsilon}_2 + |\mathbf{r}_w|\bar{A}_3\boldsymbol{\varepsilon}_3. \tag{2.15}$$

These considerations should exhibit quite clearly the analytical reason for the appearance of the scale factors. Again (2.15) can be used to evaluate $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ in terms of $A_1$, $A_2$, $A_3$; however, since (2.5) need not be valid for the arbitrary coordinate system under consideration, scalar multiplication of (2.15) by $\boldsymbol{\varepsilon}_1$, $\boldsymbol{\varepsilon}_2$, $\boldsymbol{\varepsilon}_3$ will yield *a system of three linear equations for* $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$, *the solution of which will again be of the form* (2.8).

Moreover, it follows from (2.15) that the length $|\mathbf{A}|$ of **A** in terms of the components $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ is given by

$$|\mathbf{A}|^2 = \mathbf{A} \cdot \mathbf{A} = |\mathbf{r}_u|^2(\bar{A}_1)^2 + |\mathbf{r}_v|^2(\bar{A}_2)^2 + |\mathbf{r}_w|^2(\bar{A}_3)^2 + 2|\mathbf{r}_u|\,|\mathbf{r}_v|\bar{A}_1\bar{A}_2(\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2)$$
$$+ 2|\mathbf{r}_v|\,|\mathbf{r}_w|\bar{A}_2\bar{A}_3(\boldsymbol{\varepsilon}_2 \cdot \boldsymbol{\varepsilon}_3) + 2|\mathbf{r}_w|\,|\mathbf{r}_u|\bar{A}_3\bar{A}_1(\boldsymbol{\varepsilon}_3 \cdot \boldsymbol{\varepsilon}_1).$$

But from (2.12) it follows that

$$\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2 = \frac{\mathbf{r}_u \cdot \mathbf{r}_v}{|\mathbf{r}_u|\,|\mathbf{r}_v|},$$

with two similar relations, and hence we conclude that

$$|\mathbf{A}|^2 = |\mathbf{r}_u|^2(\bar{A}_1)^2 + |\mathbf{r}_v|^2(\bar{A}_2)^2 + |\mathbf{r}_w|^2(\bar{A}_3)^2$$
$$+ 2(\mathbf{r}_u \cdot \mathbf{r}_v)\bar{A}_1\bar{A}_2 + 2(\mathbf{r}_v \cdot \mathbf{r}_w)\bar{A}_2\bar{A}_3 + 2(\mathbf{r}_w \cdot \mathbf{r}_u)\bar{A}_3\bar{A}_1. \tag{2.16}$$

Thus, in an arbitrary curvilinear coordinate system in $E_3$ *the square of the length of a vector* **A** *at a point $P$ of $E_3$ is a quadratic form in the components* $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ *of* **A**, *the coefficients of this quadratic form being functions of the positional coordinates of the point $P$ of application of* **A**.

For the case of a spherical polar coordinate system with $u = \rho$, $v = \theta$, $w = \phi$, the transformation equations (2.11) reduce to (2.1), from which it is evident that the components of $\mathbf{r}_\rho$, $\mathbf{r}_\theta$, $\mathbf{r}_\phi$ are respectively given by

$$\left.\begin{array}{l} \mathbf{r}_\rho = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \\ \mathbf{r}_\theta = (\rho\cos\theta\cos\phi, \rho\cos\theta\sin\phi, -\rho\sin\theta), \\ \mathbf{r}_\phi = (-\rho\sin\theta\sin\phi, \rho\sin\theta\cos\phi, 0), \end{array}\right\} \tag{2.17}$$

so that

$$|\mathbf{r}_\rho|^2 = 1, \qquad |\mathbf{r}_\theta|^2 = \rho^2, \qquad |\mathbf{r}_\phi|^2 = \rho^2\sin^2\theta. \tag{2.18}$$

Thus the scale factors for our spherical polar coordinate system are 1, $\rho$, $\rho\sin\theta$, which is in agreement with our previous analysis. Also, it is easily verified analytically that

$$\mathbf{r}_\rho \cdot \mathbf{r}_\theta = 0, \qquad \mathbf{r}_\theta \cdot \mathbf{r}_\phi = 0, \qquad \mathbf{r}_\phi \cdot \mathbf{r}_\rho = 0, \tag{2.19}$$

and hence the substitution of (2.18) and (2.19) in (2.15) and (2.16) respectively yields the relations (2.4) and (2.9) as special cases.

## 1.3 SOME ELEMENTARY PROPERTIES OF DETERMINANTS

In the sequel certain properties of determinants will be used repeatedly. These are listed here for future reference; it should be emphasized, however, that no attempt is made to provide a comprehensive description of the theory of determinants, (see, e.g., Hadley [1]).

Let us suppose that we are given $n^2$ quantities arranged as follows in matrix form:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{3.1}$$

For the sake of brevity we denote this matrix by $(a_{jh})$, it being understood that the indices $j, h, k, \ldots$ range over the values $1, 2, \ldots, n$ throughout this section.

The $(n \times n)$ *determinant* of the matrix (3.1) is defined to be

$$a = \sum_{j_1 \cdots j_n} (-1)^\kappa a_{1j_1} a_{2j_2} \cdots a_{nj_n}, \tag{3.2}$$

in which the summation is extended over all $n!$ permutations $j_1, j_2, \ldots, j_n$ of the first $n$ integers, and where $\kappa = 0$ or $1$ according as the permutation at hand is even or odd. The following notation is used frequently below:

$$a = \det(a_{jh}) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{3.3}$$

It is immediately obvious from the definition (3.2) that an interchange of two rows (or two columns) changes the sign of the determinant; in particular, if two rows (or two columns) are identical, the value of the determinant is zero.

Let $k$ be a fixed integer, with $1 \le k \le n$. Clearly each term of the sum on the right-hand side of (3.2) contains precisely one factor of the type $a_{kj}$, that is, one element from the $k$th row in (3.3). Hence the terms in (3.2) consist of $n$ types, the first of which possesses the factor $a_{k1}$, the second possesses the factor $a_{k2}, \ldots$, and the $n$th possesses the factor $a_{kn}$. Thus we can write (3.2) in the form

$$a = a_{k1}A_{1k} + a_{k2}A_{2k} + \cdots + a_{kn}A_{nk} = \sum_{j=1}^{n} a_{kj}A_{jk}, \tag{3.4}$$

in which the coefficient $A_{jk}$ of the factor $a_{kj}$ is called the *cofactor* of $a_{kj}$. It is a direct consequence of the definition (3.2)—as may be verified quite easily—that $A_{jk}$ is identical with the value of the $(n-1) \times (n-1)$ determinant obtained from (3.3) by the deletion of the $k$th row and the $j$th column multiplied by $(-1)^{k+j}$. Moreover, since by construction the cofactors $A_{k1}, A_{k2}, \ldots, A_{kn}$ do not contain any elements of the form $a_{kj}$, it follows from (3.4) that

$$A_{jk} = \frac{\partial a}{\partial a_{kj}}. \tag{3.5}$$

The identity (3.4) represents the expansion of a determinant in terms of the elements of its $k$th row; a similar formula may obviously be obtained for the expansion of the determinant in terms of the elements of its $k$th column:

$$a = a_{1k}A_{k1} + a_{2k}A_{k2} + \cdots + a_{nk}A_{kn} = \sum_{j=1}^{n} A_{kj}a_{jk}. \tag{3.6}$$

This suggests that we consider a sum of the type $\sum_{j=1}^{n} a_{kj}A_{jh}$ in which $h \ne k$. From the definition of the cofactor it is immediately evident that this represents the expansion of a determinant with two identical rows, namely, the $k$th and the $h$th rows, and accordingly this sum vanishes. This, together with (3.4), indicates that

$$\sum_{j=1}^{n} a_{kj}A_{jh} \begin{cases} = a & \text{if} & k = h, \\ = 0 & \text{if} & k \ne h. \end{cases} \tag{3.7}$$

This result can be written in a more compact form in terms of the so-called *Kronecker delta* $\delta_{kh}$. The latter is defined as follows:

$$\delta_{kh} \begin{cases} = 1 & \text{if} & k = h & (h, k, \ldots = 1, \ldots, n), \\ = 0 & \text{if} & k \ne h & (h, k, \ldots = 1, \ldots, n). \end{cases}$$

Thus (3.7) can be expressed as

$$\sum_{j=1}^{n} a_{kj}A_{jh} = a\delta_{kh}, \tag{3.8}$$

while (3.6) leads to an analogous relation, namely,

$$\sum_{j=1}^{n} A_{kj}a_{jh} = a\delta_{kh}. \tag{3.9}$$

Thus, provided that the determinant $a$ does not vanish, we can always construct the elements $b_{jh}$ of the matrix which is *inverse* to (3.1) by putting

$$b_{jh} = a^{-1}A_{jh}, \tag{3.10}$$

since, by virtue of (3.8) and (3.9), these quantities satisfy the conditions

$$\sum_{j=1}^{n} a_{kj}b_{jh} = \sum_{j=1}^{n} b_{kj}a_{jh} = \delta_{kh}. \tag{3.11}$$

Let us now consider two $(n \times n)$ matrices $(a_{jh})$, $(c_{jh})$, whose determinants are respectively denoted by $a$ and $c$. The matrix whose elements are defined by

$$f_{kh} = \sum_{j=1}^{n} a_{kj}c_{jh}, \tag{3.12}$$

is the product of the matrices $(a_{jh})$, $(c_{jh})$ (in this order); it is easily verified that the value $f$ of the determinant of this product matrix is given by

$$f = ac. \tag{3.13}$$

In particular, the sums which appear on the left-hand side of (3.11) represent the elements of the product of the matrix $(a_{jh})$ and its inverse $(b_{jh})$.

In many of the applications to be considered below, the elements $a_{jh}$ are not constants, but are in fact functions of one or more variables, and it is frequently necessary to evaluate the derivatives of the corresponding determinants. Let us suppose, then, that the $a_{kj}$ are continuously differentiable functions of $m$ parameters $t^\alpha$ $(\alpha = 1, \ldots, m)$:

$$a_{kj} = a_{kj}(t^1, \ldots, t^m).$$

Clearly the partial derivatives of the determinant $a$ are given by

$$\frac{\partial a}{\partial t^\alpha} = \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\partial a}{\partial a_{kj}} \frac{\partial a_{kj}}{\partial t^\alpha}. \tag{3.14}$$

However, by means of (3.5) this can be written in terms of the cofactors:

$$\frac{\partial a}{\partial t^\alpha} = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{jk} \frac{\partial a_{kj}}{\partial t^\alpha}. \tag{3.15}$$

This simple formula will be found to be extremely useful.

## PROBLEMS

1.1    In $E_3$ the quantity $\varepsilon_{jhk}$ is defined by

$$\varepsilon_{jhk} = \begin{cases} +1 \text{ if } (j, h, k) \text{ is an even permutation of } (1, 2, 3), \\ -1 \text{ if } (j, h, k) \text{ is an odd permutation of } (1, 2, 3), \\ 0 \text{ if } (j, h, k) \text{ has any other set of values.} \end{cases}$$

If $C_{hk}$ and $C_j$ are defined by (1.15) and (1.17), respectively, show that

$$C_j = \frac{1}{2} \sum_{h=1}^{3} \sum_{k=1}^{3} \varepsilon_{jhk} C_{hk} \qquad (j = 1, 2, 3).$$

1.2    If the $3 \times 3$ matrix $(a_{ij})$ has determinant $a$, show that

$$a\varepsilon_{ijk} = \sum_{h=1}^{3} \sum_{l=1}^{3} \sum_{m=1}^{3} a_{ih} a_{jl} a_{km} \varepsilon_{hlm},$$

where $\varepsilon_{ijk}$ is defined in Problem 1.1.

1.3    Show that

$$\sum_{j=1}^{n} \delta_{jj} = n,$$

$$\sum_{j=1}^{n} \delta_{ij} \delta_{jk} = \delta_{ik} \qquad (i, k = 1, \ldots, n),$$

and

$$\sum_{j=1}^{n} \delta_{ij} A_j = A_i \qquad (i = 1, \ldots, n),$$

where $A_1, \ldots, A_n$ are arbitrary quantities.

1.4    If $B_1, \ldots, B_n$ are arbitrary quantities and $a_{ij} = B_i B_j$ $(i, j = 1, \ldots, n)$ prove that $\det(a_{ij}) = 0$.

1.5    If $A_{jk}$ is the cofactor of $a_{kj}$ show that

$$\det(A_{jk}) = a^{n-1}$$

where $a = \det(a_{kj})$.

1.6    If $a_{jk} = -a_{kj}$ for $j, k = 1, \ldots, n$ prove that, for $n$ odd,

$$\det(a_{jk}) = 0.$$

1.7    If an elastic medium is isotropic then Hooke's law assumes the form

$$\tau_{jk} = \lambda \sum_{i=1}^{3} s_{ii} \delta_{jk} + 2\mu s_{jk} \qquad (j, k = 1, 2, 3),$$

where $\lambda, \mu$ are constants. Show that in this case if $c_{jklm} = c_{jkml}$, then the relation (1.7) implies that

$$c_{jklm} = \lambda \delta_{jk} \delta_{lm} + \mu(\delta_{jl} \delta_{km} + \delta_{jm} \delta_{kl}).$$

Hence show that

$$c_{jklm} = c_{kjlm}, \qquad c_{jklm} = c_{lmjk}.$$

Establish that $c_{jklm} + c_{jmkl} + c_{jlmk} = 0$ if and only if $\lambda = -2\mu$.

# 2

# AFFINE TENSOR ALGEBRA IN EUCLIDEAN GEOMETRY

Since the tensor calculus is concerned with the behavior of various entities under the transition from a given coordinate system to another, we begin with a detailed analysis of the simplest type of coordinate transformation, namely, the orthogonal transformation in a three-dimensional Euclidean space $E_3$. The resulting transformation properties of the components of vectors suggest the introduction of a special class of tensors, the so-called affine tensors, whose definition is formulated in terms of the behavior of their components under orthogonal transformations in $E_3$. This gives rise to a simple algebraic theory of tensors which may be regarded as an immediate extension of elementary vector algebra. Moreover, the basic definitions are such as to allow for an effortless transition from three to $n$ dimensions. However, by definition, orthogonal transformations are linear, and consequently the theory of affine tensors is far too restricted for most purposes. For instance, whenever curvilinear coordinates are introduced, one is confronted with nonlinear coordinate transformations. By abstracting some of the salient features of affine tensor algebra, however, one may formulate in a natural manner a more general definition of the tensor concept which is no longer dependent on the linearity inherent in orthogonal transformations. An immediate application of these ideas is the study of parallel vector fields in $E_n$ referred to curvilinear coordinates; this investigation is of relevance also to the more general theory developed in subsequent chapters.

## 2.1 ORTHOGONAL TRANSFORMATIONS IN $E_3$

In this section we are concerned with the transformation from a given rectangular coordinate system to another, it being assumed that these systems have a common origin at a point $O$ of our three-dimensional Euclidean space $E_3$. The unit vectors in the directions of the positive $Ox^1$-, $Ox^2$-, and $Ox^3$-axes are again denoted by $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. Since these vectors are mutually orthogonal we may write

$$\mathbf{e}_h \cdot \mathbf{e}_j = \delta_{hj} \qquad (h, j = 1, 2, 3), \tag{1.1}$$

in which $\delta_{jh}$ denotes the Kronecker delta as defined in Section 1.3. The vectors $\mathbf{e}_j$ are said to constitute an *orthonormal basis* $\{\mathbf{e}_j\}$ of $E_3$; any vector of $E_3$ can be expressed as a linear combination of these basis vectors (see, e.g., Hadley [1]). In particular, for the basis vectors of another orthonormal basis centered at $O$, denoted provisionally by $\mathbf{f}_j$ ($j = 1, 2, 3$), one has a representation of the form

$$\mathbf{f}_1 = a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + a_{13}\mathbf{e}_3,$$
$$\mathbf{f}_2 = a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + a_{23}\mathbf{e}_3,$$
$$\mathbf{f}_3 = a_{31}\mathbf{e}_1 + a_{32}\mathbf{e}_2 + a_{33}\mathbf{e}_3,$$

in which the $a_{jh}$ are real parameters which depend on the directions of the vectors $\mathbf{f}_j$ relative to the basis $\{\mathbf{e}_j\}$. This representation can be written more compactly as

$$\mathbf{f}_j = \sum_{h=1}^{3} a_{jh}\mathbf{e}_h. \tag{1.2}$$

Clearly this transformation is characterized completely by the $3 \times 3$ matrix $(a_{jh})$, whose elements possess a simple geometrical interpretation: when the scalar products of (1.2) with $\mathbf{e}_k$ ($k = 1, 2, 3$) are formed, the relations (1.1) being taken into account, it is seen that

$$\mathbf{f}_j \cdot \mathbf{e}_k = \sum_{h=1}^{3} a_{jh}(\mathbf{e}_h \cdot \mathbf{e}_k) = \sum_{h=1}^{3} a_{jh}\delta_{hk} = a_{jk}, \tag{1.3}$$

or, since $\mathbf{f}_j$, $\mathbf{e}_k$ are unit vectors;

$$a_{jk} = \cos(\mathbf{f}_j, \mathbf{e}_k) \qquad (j, k = 1, 2, 3). \tag{1.4}$$

Thus the coefficients $a_{jk}$ in (1.2) are simply the cosines of the angles between the various basis vectors of the two orthonormal systems, which in turn implies that the $a_{jk}$ cannot assume arbitrary values or be independent of each other. Indeed, there exists a set of relations between them which may be easily obtained as follows.

For the orthonormal basis $\{\mathbf{f}_j\}$ we have, as in (1.1),

$$\mathbf{f}_j \cdot \mathbf{f}_k = \delta_{jk} \qquad (j, k = 1, 2, 3),$$

so that by (1.2)

$$\delta_{jk} = \left(\sum_{h=1}^{3} a_{jh}\mathbf{e}_h\right) \cdot \left(\sum_{l=1}^{3} a_{kl}\mathbf{e}_l\right) = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh}a_{kl}(\mathbf{e}_h \cdot \mathbf{e}_l),$$

or, if we invoke (1.1),

$$\delta_{jk} = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh}a_{kl}\,\delta_{hl} = \sum_{h=1}^{3} a_{jh}a_{kh}. \tag{1.5}$$

The transpose $a_{jl}$ of the elements $a_{lj}$ of the matrix $(a_{lj})$ is to be denoted by $a_{lj}^*$, and accordingly (1.5) can be written in the form

$$\sum_{h=1}^{3} a_{jh}a_{hk}^* = \delta_{jk} \qquad (j, k = 1, 2, 3). \tag{1.6}$$

These relations represent a necessary condition which the coefficients $a_{jh}$ in (1.2) must satisfy in order that (1.2) represent a transformation from a given orthonormal basis to another. By reversing the argument it is easily shown that (1.6) is also a sufficient condition. Any linear transformation which satisfies (1.6) is called an *orthogonal* transformation and the corresponding matrix $(a_{jh})$ is said to be orthogonal.

It is evident from the definition given in Section 1.3 that $\det(a_{jh}^*) = \det(a_{hj})$; it therefore follows from (1.6), (1.3.12), and (1.3.13) that

$$\det(a_{jh})\det(a_{hk}^*) = a^2 = \det(\delta_{jk}) = 1,$$

so that the determinant of any orthogonal matrix satisfies the condition

$$a = \pm 1. \tag{1.7}$$

Thus $a \neq 0$, and the inverse of the matrix $(a_{ij})$ exists. In fact, a comparison of (1.3.11) with (1.6) indicates that *the inverse of an orthogonal matrix is identical with its transpose.* This state of affairs may be illustrated geometrically by the fact that the matrix of the inverse of the transformation (1.2) is, in fact, the transpose of the matrix $(a_{jh})$ which characterizes (1.2). For, if we multiply (1.2) by $a_{kj}^*$, after which we sum over $j$ from 1 to 3, at the same time taking (1.6) into account, we find that

$$\sum_{j=1}^{3} a_{kj}^*\mathbf{f}_j = \sum_{j=1}^{3} \sum_{h=1}^{3} a_{kj}^* a_{jh}\mathbf{e}_h = \sum_{h=1}^{3} \delta_{kh}\mathbf{e}_h,$$

or

$$\mathbf{e}_k = \sum_{j=1}^{3} a_{kj}^*\mathbf{f}_j. \tag{1.8}$$

This relation expresses the basis vectors $\mathbf{e}_k$ in terms of $\mathbf{f}_j$ and accordingly represents the inverse of the original transformation (1.2). Moreover, in the same manner in which (1.2) gives rise to (1.4), it is seen that (1.8) implies that

$$a_{kj}^* = \cos(\mathbf{e}_k, \mathbf{f}_j), \tag{1.9}$$

which is obviously consistent with the definition of $a_{kj}^*$.

Let $\{\mathbf{g}_l\}$ denote the basis vectors of a third orthonormal basis whose origin is located at $O$. Clearly each vector $\mathbf{g}_l$ can be represented in terms of either one of the original bases: for instance,

$$\mathbf{g}_l = \sum_{j=1}^{3} c_{lj}\mathbf{f}_j, \tag{1.10}$$

in which we substitute from (1.2), so that

$$\mathbf{g}_l = \sum_{h=1}^{3} \sum_{j=1}^{3} c_{lj}a_{jh}\mathbf{e}_h. \tag{1.11}$$

Thus if we write

$$\mathbf{g}_l = \sum_{h=1}^{3} p_{lh}\mathbf{e}_h, \tag{1.12}$$

it follows from (1.11) that the corresponding matrix elements $p_{lh}$ are given by

$$p_{lh} = \sum_{j=1}^{3} c_{lj}a_{jh}, \tag{1.13}$$

which clearly demonstrates that the matrix $(p_{lh})$ is the product of the matrices $(c_{lh})$ and $(a_{jh})$.

The transformation characterized by the matrix $(p_{lh})$ is said to be the *product* of the transformations (1.10) and (1.2). Furthermore, since $(p_{lh})$ is orthogonal by construction, it follows that the product of any two orthogonal transformations is again orthogonal. Also, since matrix products obviously satisfy an associative law, we may infer that the set of all orthogonal transformations in $E_3$ is endowed with an associative binary operation, namely, the product, which is such that the product of any two elements is again an element of the set. Moreover, this set contains a unit element, defined as the identity transformation whose coefficients are merely the Kronecker deltas, while the existence of the inverse of each element is assured. Accordingly the set of all orthogonal transformations in $E_3$ forms a group, the so-called *orthogonal group*, which is usually denoted by $O_3$.

## 2.2 TRANSFORMATION PROPERTIES OF AFFINE VECTOR COMPONENTS AND RELATED CONCEPTS

Let us consider the position vector $\mathbf{r}$ of a point $P$ of our three-dimensional Euclidean space $E_3$ relative to the common origin $O$ of two orthonormal bases with basis vectors $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_j\}$, respectively.

The coordinates of $P$ in the rectangular systems defined by these bases are provisionally denoted by $x^h$ and $y^j$, respectively ($h, j = 1, 2, 3$), which implies that

$$\mathbf{r} = \sum_{h=1}^{3} x^h \mathbf{e}_h, \tag{2.1}$$

and

$$\mathbf{r} = \sum_{j=1}^{3} y^j \mathbf{f}_j. \tag{2.2}$$

Clearly there must exist a relationship between the two sets $\{x^h\}$ and $\{y^j\}$; this is easily found explicitly by the application of (1.8) to the right-hand side of (2.1), which gives

$$\mathbf{r} = \sum_{h=1}^{3} \sum_{j=1}^{3} x^h a_{hj}^* \mathbf{f}_j = \sum_{j=1}^{3} \sum_{h=1}^{3} a_{jh} x^h \mathbf{f}_j.$$

This result is compared with the right-hand side of (2.2), it being noted that the coefficients of each $\mathbf{f}_j$ in these relations must necessarily be identical, so that

$$y^j = \sum_{h=1}^{3} a_{jh} x^h \qquad (j = 1, 2, 3). \tag{2.3}$$

This is the transformation formula by means of which the $y$-coordinates of any point $P$ of $E_3$ are given in terms of the original $x$-coordinates. It is not surprising that the coefficients $a_{jh}$ which completely specify this transformation are precisely those which appear in the transition (1.2) from the orthonormal basis $\{\mathbf{e}_j\}$ to the basis $\{\mathbf{f}_j\}$.

The inverse of (2.3), which corresponds to (1.8), is therefore given by

$$x^k = \sum_{j=1}^{3} a_{kj}^* y^j = \sum_{j=1}^{3} a_{jk} y^j \qquad (k = 1, 2, 3). \tag{2.4}$$

The importance of the formulae (2.3) and (2.4) is due to the fact that they will enable us to specify the transformation of any given function of the coordinates under a change of orthonormal bases. Examples of this state of affairs are encountered below.

At this stage it is perhaps appropriate to introduce a change of notation, which allows us to write our subsequent analysis in a more compact form. Beginning with a given orthonormal system $\{\mathbf{e}_j\}$, the "transformed" basis $\{\mathbf{f}_j\}$ resulting from the application of the transformation (1.2) will be denoted by $\{\bar{\mathbf{e}}_j\}$, and accordingly the coordinates $y^j$ of a point $P$ of $E_3$ relative to the latter system will be represented by $\bar{x}^j$. Thus we must now rewrite (1.2) and (1.8), respectively, as

$$\bar{\mathbf{e}}_j = \sum_{h=1}^{3} a_{jh} \mathbf{e}_h \tag{2.5}$$

and

$$\mathbf{e}_k = \sum_{j=1}^{3} a_{kj}^* \bar{\mathbf{e}}_j = \sum_{j=1}^{3} a_{jk} \bar{\mathbf{e}}_j. \tag{2.6}$$

Similarly the transformation equations (2.3) and (2.4) now assume the forms

$$\bar{x}^j = \sum_{h=1}^{3} a_{jh} x^h \qquad (j = 1, 2, 3), \tag{2.7}$$

and

$$x^k = \sum_{j=1}^{3} a_{jk} \bar{x}^j \qquad (k = 1, 2, 3), \tag{2.8}$$

respectively. We shall adhere to this notation throughout.

Thus far we have merely discussed the transformation of coordinates under a change of basis. It is now necessary to consider the transformation properties of vectors, or more precisely, the transformation properties of the components of vectors. As long as we are dealing with rectangular coordinate systems—as is indeed the case here—there are no difficulties. Relative to a given orthonormal basis $\{\mathbf{e}_j\}$ we can represent any vector $\mathbf{A}$ in the form (1.11); if we apply (2.6) to the right-hand side of (1.11), we obtain

$$\mathbf{A} = \sum_{h=1}^{3} \sum_{j=1}^{3} a_{jh} A_h \bar{\mathbf{e}}_j. \tag{2.9}$$

However, relative to the new orthonormal basis $\{\bar{\mathbf{e}}_j\}$, the vector $\mathbf{A}$ is represented as

$$\mathbf{A} = \sum_{j=1}^{3} \bar{A}_j \bar{\mathbf{e}}_j, \tag{2.10}$$

where $\bar{A}_1$, $\bar{A}_2$, $\bar{A}_3$ denote the components of $\mathbf{A}$ in the latter system. Again, the coefficients of $\bar{\mathbf{e}}_j$ on the right-hand sides of (2.9) and (2.10) must coincide,

and accordingly we conclude that

$$\bar{A}_j = \sum_{h=1}^{3} a_{jh} A_h \qquad (j = 1, 2, 3). \tag{2.11}$$

This, then, is the required transformation law for the components of a vector $\mathbf{A}$; it is easily seen, as above, that the inverse of (2.11) is given by

$$A_k = \sum_{j=1}^{3} a_{jk} \bar{A}_j \qquad (k = 1, 2, 3). \tag{2.12}$$

The simple argument leading to this conclusion depends crucially on the existence of a rectangular coordinate system and on the conceptual interpretation in terms of (1.1.1) of a vector as a directed segment which is completely determined by the oriented lengths of the projections of the latter onto the coordinate axes. However, we have seen that, in the case of curvilinear coordinate systems, this approach to the vector concept possesses no immediate counterpart.

Thus even while we remain temporarily within the context of orthogonal transformations as applied to orthonormal bases, it would seem advisable to abandon the earlier, more intuitive point of view concerning the idea of a vector, and to redefine this concept in terms of certain abstractions.

Our first step in this direction is achieved by the formulation of the following.

**DEFINITION**

*A set of three quantities $(A_1, A_2, A_3)$ is said to constitute the components of an affine vector $\mathbf{A}$, if, under the orthogonal transformation (2.7) of rectangular coordinates in $E_3$, the transformed quantities $\bar{A}_1, \bar{A}_2, \bar{A}_3$ are given by formula (2.11).*

*Remark 1.* Some texts use the nomenclature *affine orthogonal vector*, or *Cartesian vector* for the concept defined here. Whereas the terminology per se is immaterial, it is essential that the adjectives *affine*, or *orthogonal*, or *Cartesian*, be appended in this context, for they emphasize the fact that the definition depends vitally on the stipulation that the transformations concerned be *orthogonal*. [At a later stage more general definitions of the vector concept will be formulated which are not subject to restrictions of this kind, and which are therefore suitable also for curvilinear coordinate systems.]

*Remark 2.* The vector $\mathbf{A}$ thus defined should be regarded as an entirely new entity superimposed on the geometry of $E_3$. It is *not* an element of $E_3$: the elements of the latter are merely its points. [Very often no clear distinction

is made between a point $P$ of $E_3$, and the position vector $\mathbf{r}$ of $P$, and accordingly vectors of $E_3$ are frequently treated as elements of $E_3$. This misconception is enhanced by the fact that, under the special circumstances treated here, the transformation (2.7) of point-coordinates is formally identical with the transformation (2.11) of vector components; however, it will soon be evident that under slightly more general conditions this similarity disappears in toto.]

*Remark 3.* It should be noted that the above definition of an affine vector is entirely independent of the concept of the length of the vector. Indeed, the latter is introduced by means of a secondary definition, which is meaningful solely under the special circumstances treated here. [Under more general conditions it may not even be feasible to define the length of a vector.]

The *length* $|\mathbf{A}|$ of a vector $\mathbf{A}$, whose components relative to the orthonormal basis $\{\mathbf{e}_j\}$ are denoted by $A_1, A_2, A_3$, is defined by

$$|\mathbf{A}|^2 = \sum_{k=1}^{3} (A_k)^2, \tag{2.13}$$

it being understood that $|\mathbf{A}|$ refers to the positive square root of the right-hand side above.

This definition involves specific reference to a given orthonormal basis: it is easily seen, however, that the numerical value of $|\mathbf{A}|$ as defined by (2.13) is independent of the actual choice of basis. For, if we substitute in (2.13) from (2.12), we find that

$$|\mathbf{A}|^2 = \sum_{k=1}^{3} \sum_{j=1}^{3} \sum_{h=1}^{3} a_{jk} a_{hk} \bar{A}_j \bar{A}_h, \tag{2.14}$$

while according to (1.6)

$$\sum_{k=1}^{3} a_{jk} a_{hk} = \delta_{jh}, \tag{2.15}$$

so that (2.14) becomes

$$|\mathbf{A}|^2 = \sum_{j=1}^{3} \sum_{h=1}^{3} \delta_{jh} \bar{A}_j \bar{A}_h = \sum_{h=1}^{3} (\bar{A}_h)^2. \tag{2.16}$$

But the right-hand side of this relation is precisely the expression for $|\mathbf{A}|^2$ which would have appeared if the definition (2.13) had been written in terms of the orthonormal basis defined by $\{\bar{\mathbf{e}}_j\}$. Combining (2.13) and (2.16) we see that

$$|\mathbf{A}|^2 = \sum_{k=1}^{3} (A_k)^2 = \sum_{k=1}^{3} (\bar{A}_k)^2, \tag{2.17}$$

a phenomenon which is often referred to as the *invariance of length under orthogonal transformations*. This also justifies our definition (2.13): clearly one would not want the concept of the length of a vector to be dependent on the choice of the rectangular coordinate system. It should be stressed, however, that this invariance depends on the fact that we are dealing solely with orthogonal transformations in the present context. Thus under the more general circumstances to be discussed below, the definition (2.13) ceases to be appropriate. This, of course, is not surprising in view of our observations in Section 1.2 concerning curvilinear coordinate systems.

*Remark 4.* For future reference, we indicate how the transformation equations (2.11) and (2.12) can be written in a somewhat different form. Let us differentiate the equations (2.7), which represent a point transformation, with respect to $x^h$, regarding $x^1$, $x^2$, $x^3$ as independent variables. This gives

$$\frac{\partial \bar{x}^j}{\partial x^h} = a_{jh}, \tag{2.18}$$

while the corresponding inverse is similarly obtained from (2.8):

$$\frac{\partial x^k}{\partial \bar{x}^j} = a_{jk}. \tag{2.19}$$

When this is substituted in (2.11) and (2.12) we see that the latter become

$$\bar{A}_j = \sum_{h=1}^{3} \frac{\partial \bar{x}^j}{\partial x^h} A_h \tag{2.20}$$

and

$$A_k = \sum_{j=1}^{3} \frac{\partial x^k}{\partial \bar{x}^j} \bar{A}_j, \tag{2.21}$$

respectively.

In Section 1.1 we encountered certain quantities whose components—in contrast to vectors—require for their specification more than one index. We now investigate the transformation properties of these entities.

The simplest example of this kind is furnished by the products of the components of two vectors **A** and **B**. The transforms under (2.7) of the components of **A** are given by (2.11); we have similarly for the transforms of the components of **B**:

$$\bar{B}_k = \sum_{l=1}^{3} a_{kl} B_l, \tag{2.22}$$

so that

$$\bar{A}_j \bar{B}_k = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} A_h B_l. \tag{2.23}$$

We can regard the nine products $A_h B_l$ ($h = 1, 2, 3$; $l = 1, 2, 3$) as the entries of a $3 \times 3$ matrix, and (2.23) then gives us the transformation law of these matrix elements. It will be seen that (2.23) is characteristic of the transformation properties of certain types of tensors; however, prior to the formulation of the precise definition of this concept, we examine some less trivial special cases.

If we interchange the order of the indices $j$ and $k$ in (2.23), we have

$$\bar{A}_k \bar{B}_j = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{kh} a_{jl} A_h B_l = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jl} a_{kh} A_h B_l.$$

In the expression on the right one may replace the index $h$ by $l$, and similarly $l$ by $h$, so that

$$\bar{A}_k \bar{B}_j = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} A_l B_h.$$

This is subtracted from (2.23), which yields

$$\bar{A}_j \bar{B}_k - \bar{A}_k \bar{B}_j = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} (A_h B_l - A_l B_h).$$

Thus if we define the components $C_{hl}$ of the vector product **A** × **B** in accordance with (1.1.15), namely, by

$$C_{hl} = A_h B_l - A_l B_h, \tag{2.24}$$

we see that the transformed components of **A** × **B** are given by

$$\bar{C}_{jk} = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} C_{hl}. \tag{2.25}$$

It should be observed that the quantities (2.24) obey a transformation law which is formally identical with that exemplified by (2.23).

Another example of a special kind is provided by the Kronecker deltas. These quantities assume the numerical values 0 and 1 irrespective of the choice of the coordinate system, so that we are entitled to write also in the $\bar{x}$-coordinate system: $\bar{\delta}_{jk} = 1$ if $j = k$, while $\bar{\delta}_{jk} = 0$ if $j \neq k$. Thus in this particular instance we have $\bar{\delta}_{jk} = \delta_{jk}$. On the other hand, we may then infer from (1.5) that

$$\bar{\delta}_{jk} = \sum_{h=1}^{3} a_{jh} a_{kh} = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} \delta_{hl}, \tag{2.26}$$

from which it is evident that the Kronecker deltas are also quantities which satisfy transformation laws of the type (2.23).

Let us now turn to the inertia tensor as defined in Section 1.1. If we use the Kronecker delta once more we can write the expressions (1.1.10) and (1.1.11) for the components of the inertia tensor relative to our $x$-coordinate system in the form

$$I_{jk} = \int \rho i_{jk} \, dV, \tag{2.27}$$

in which we have put

$$i_{jk} = \left( \sum_{m=1}^{3} x^m x^m \right) \delta_{jk} - x^j x^k. \tag{2.28}$$

When we subject these quantities to the coordinate transformation (2.7), we note that the sum $\sum_{m=1}^{3} x^m x^m$ is invariant: this is a special case of the identity (2.17) with $A_j = x^j$. Thus substituting directly from (2.26) and (2.7), we find that

$$
\begin{aligned}
\bar{i}_{jk} &= \sum_{m=1}^{3} (\bar{x}^m \bar{x}^m) \bar{\delta}_{jk} - \bar{x}^j \bar{x}^k \\
&= \sum_{m=1}^{3} (x^m x^m) \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} \delta_{hl} - \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} x^h x^l \\
&= \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} \left[ \sum_{m=1}^{3} (x^m x^m) \delta_{hl} - x^h x^l \right] \\
&= \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} i_{hl},
\end{aligned}
\tag{2.29}
$$

where, in the last step, we have used the definition (2.28) once more. Thus the quantities $i_{jk}$ transform according to a law which is analogous to (2.23). In order to determine in which manner the components (2.27) transform, we have to integrate (2.29) over the rigid body under consideration. To this end we note that the coefficients $a_{jh}$ in (2.29) are independent of the positional coordinates; thus they can be treated as constants for the purposes of integration. Also, the density $\rho$ is obviously unaffected by the coordinate transformation (2.7). However, for any volume integral referred to rectangular axes we have from the usual rule for the transformation of integrals:

$$\int dV = \int dx^1 \, dx^2 \, dx^3 = \int \left| \frac{\partial(x^1, x^2, x^3)}{\partial(\bar{x}^1, \bar{x}^2, \bar{x}^3)} \right| d\bar{x}^1 \, d\bar{x}^2 \, d\bar{x}^3,$$

in which the Jacobian is the determinant of the partial derivatives $\partial x^j / \partial \bar{x}^h$,

which, according to (2.19) and (1.7), is equal to $\det(a_{jh}) = a = \pm 1$. Thus the volume element $dV$ is unaffected by the transformation, and the integration of (2.29) therefore yields

$$\bar{I}_{jk} = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} I_{hl}. \tag{2.30}$$

Again we observe that the components of the inertia tensor transform in accordance with a transformation law of the type (2.23).

Finally, let us derive the transformation properties of the so-called stress tensor of an elastic body as defined in Section 1.1. Again we construct a rectangular coordinate system with origin at a fixed point $O$ of the body, the corresponding basis vectors being denoted by $\{\mathbf{e}_j\}$. We must now attempt to obtain the components of the stress tensor relative to another rectangular coordinate system with origin at $O$, whose basis vectors are $\{\bar{\mathbf{e}}_l\}$. To this end we consider a small tetrahedron, whose sides are portions of the planes $x^1 = 0, x^2 = 0, x^3 = 0$, and $\bar{x}^l = $ const. (for some fixed value of $l$), the areas of these portions being respectively denoted by $dS_{(1)}, dS_{(2)}, dS_{(3)},$ and $d\bar{S}_{(l)}$ (see Figure 4). Since the latter is, by definition, intersected orthogonally
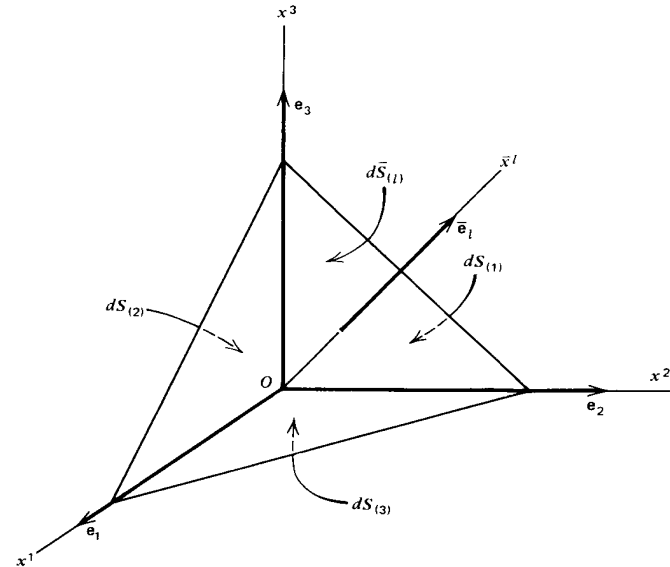


Fig. 4

by the vector $\bar{\mathbf{e}}_l$, it follows that $dS_{(k)} = d\bar{S}_{(l)} \cos(\bar{\mathbf{e}}_l, \mathbf{e}_k)$ $(k = 1, 2, 3)$, or because of (1.4),

$$dS_{(k)} = a_{lk} \, d\bar{S}_{(l)}. \tag{2.31}$$

Now by definition of the stress tensor components, the force per unit area across that face of the tetrahedron which is normal to $\mathbf{e}_k$ is $\mathbf{F}_{(k)} = \sum_{j=1}^{3} \tau_{kj} \mathbf{e}_j$, so that the total force across this face is

$$\mathbf{F}_{(k)} \, dS_{(k)} = \sum_{j=1}^{3} \tau_{kj} \, dS_{(k)} \mathbf{e}_j.$$

Hence the sum of the forces across the three faces $x^1 = 0$, $x^2 = 0$, $x^3 = 0$ of the tetrahedron is

$$\sum_{k=1}^{3} \sum_{j=1}^{3} \tau_{kj} \, dS_{(k)} \mathbf{e}_j = \sum_{k=1}^{3} \sum_{j=1}^{3} a_{lk} \tau_{kj} \mathbf{e}_j \, d\bar{S}_{(l)}, \tag{2.32}$$

where, in the second step, we have used (2.31). But this sum must be equal to the total force across the remaining face of the tetrahedron; again, by definition of the stress tensor, this force is

$$\sum_{h=1}^{3} \bar{\tau}_{lh} \bar{\mathbf{e}}_h \, d\bar{S}_{(l)}. \tag{2.33}$$

This is identical with (2.32), and accordingly the latter must be referred to the $\{\bar{\mathbf{e}}_h\}$ basis, which is done by substitution from (2.6), the common factor $d\bar{S}_{(l)}$ being omitted at the same time. It is thus found that

$$\sum_{h=1}^{3} \bar{\tau}_{lh} \bar{\mathbf{e}}_h = \sum_{h=1}^{3} \sum_{k=1}^{3} \sum_{j=1}^{3} a_{lk} \tau_{kj} a_{hj} \bar{\mathbf{e}}_h,$$

and hence we obtain by equating coefficients of $\bar{\mathbf{e}}_h$ on either side:

$$\bar{\tau}_{lh} = \sum_{k=1}^{3} \sum_{j=1}^{3} a_{lk} a_{hj} \tau_{kj}, \tag{2.34}$$

which is again a transformation law of the type (2.23).

The above examples provide a convincing motivation for the formulation of the following definition.

## DEFINITION

*A set of quantities $T_{jk}$ ($j = 1, 2, 3$; $k = 1, 2, 3$), is said to constitute the components of an affine tensor of rank 2, if, under the orthogonal coordinate transformation (2.7), these quantities transform according to the transformation law*

$$\bar{T}_{jk} = \sum_{h=1}^{3} \sum_{l=1}^{3} a_{jh} a_{kl} T_{hl}. \tag{2.35}$$

*Remark 1.* Again the adjective *affine* (or, as in some texts, *affine orthogonal*, or *Cartesian*) is used in order to emphasize that the coordinate transformation referred to in this definition is an orthogonal one. Furthermore, it should also be pointed out that affine tensors are *not* elements of $E_3$; they are, in fact, entirely new entities defined against the background of the theory of orthogonal transformations in $E_3$.

*Remark 2.* The definition includes the phrase "rank 2." This is done because it refers to a set of quantities whose specification depends on $3^2$ components $T_{jk}$, or equivalently, whose transformation law (2.35) involves a polynomial of order 2 in the transformation coefficients $a_{jl}$.

*Remark 3.* The inverse of (2.35) exists, and is seen to be

$$T_{hl} = \sum_{m=1}^{3} \sum_{p=1}^{3} a_{mh} a_{pl} \bar{T}_{mp}. \tag{2.36}$$

This is an immediate consequence of the fact that (2.7) is an orthogonal transformation; clearly (2.36) may be obtained directly from (2.35) by means of (1.6).

*Remark 4.* Again, with the aid of (2.18) and (2.19) the transformation laws (2.35) and (2.36) can be written in the form

$$\bar{T}_{jk} = \sum_{h=1}^{3} \sum_{l=1}^{3} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^k}{\partial x^l} T_{hl}, \tag{2.37}$$

and

$$T_{hl} = \sum_{m=1}^{3} \sum_{p=1}^{3} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^l}{\partial \bar{x}^p} \bar{T}_{mp}. \tag{2.38}$$

## 2.3   GENERAL AFFINE TENSOR ALGEBRA

Thus far we have restricted our attention to coordinate transformations in three-dimensional Euclidean spaces. This was done mainly for the sake of the physical examples studied in the previous section: from a mathematical point of view the restriction to three dimensions is not essential; indeed, it is undesirable.

Accordingly we now consider an $n$-dimensional Euclidean space $E_n$, and by means of abstractions suggested by our previous discussion, we develop the algebra of general affine tensors (see, e.g., Jeffreys [1], Synge and Schild [1], and Temple [1]).

An orthonormal system $\{\mathbf{e}_j\}$ in $E_n$ consists of $n$ mutually orthogonal unit vectors. Any other orthonormal system $\{\bar{\mathbf{e}}_j\}$ may be obtained from the

first by means of the linear transformation

$$\bar{\mathbf{e}}_j = \sum_{h=1}^{n} a_{jh}\mathbf{e}_h \qquad (j = 1, \ldots, n), \tag{3.1}$$

provided that the coefficients $a_{jh}$ satisfy the orthogonality condition

$$\delta_{jk} = \sum_{h=1}^{n} a_{jh}a_{kh} \qquad (j, k = 1, \ldots, n), \tag{3.2}$$

which, as before, implies that

$$a = \det(a_{jk}) = \pm 1. \tag{3.3}$$

Thus (3.1) is once more called an *orthogonal* transformation; if, in addition, $a = +1$, the transformation is said to be *proper orthogonal*.

The coordinates of a point $P$ relative to the orthonormal systems $\{\mathbf{e}_j\}$ and $\{\bar{\mathbf{e}}_j\}$ are respectively denoted by $x^j$ ($j = 1, \ldots, n$), and $\bar{x}^j$ ($j = 1, \ldots, n$); these coordinates are related according to the transformation equations

$$\bar{x}^j = \sum_{h=1}^{n} a_{jh}x^h \qquad (j = 1, \ldots, n), \tag{3.4}$$

whose inverse is

$$x^h = \sum_{j=1}^{n} a_{jh}\bar{x}^j \qquad (h = 1, \ldots, n). \tag{3.5}$$

We are now in a position to formulate the most general definition of an affine tensor.

## DEFINITION

*Let $r$ be any positive integer. A set of $n^r$ quantities $T_{h_1 h_2 \ldots h_r}$ is said to constitute the components of an affine tensor of rank $r$, if, under the orthogonal coordinate transformation (3.4), these quantities transform according to the transformation law*

$$\bar{T}_{j_1 j_2 \ldots j_r} = \sum_{h_1=1}^{n} \sum_{h_2=1}^{n} \cdots \sum_{h_r=1}^{n} a_{j_1 h_1} a_{j_2 h_2} \cdots a_{j_r h_r} T_{h_1 h_2 \ldots h_r}$$

$$(j_1, j_2, \ldots, j_r = 1, \ldots, n). \tag{3.6}$$

*Remark 1.* If, in this relation, we put $n = 3$, $r = 2$, and replace $j_1, j_2$ respectively by $j, k$, it is obvious that we obtain the transformation law (2.35) of an affine tensor of rank 2 in the three-dimensional case. Moreover, with $n = 3$ and $r = 1$, we obtain the transformation law (2.11) of an affine vector in three dimensions. *Thus we henceforth consider vectors to be tensors of rank* 1.

In both geometry and physics one frequently encounters quantities which depend on positional coordinates, but whose numerical values are independent of the choice of the coordinate system. Examples of this kind are represented by the lengths of vectors, or the density of a material, or the temperature of an object. Let us denote such a function of position by $\phi(x^1, \ldots, x^n)$ relative to the orthonormal basis defined by $\{\mathbf{e}_j\}$. If the value of $\phi$ is indeed independent of the choice of the coordinate system, we have

$$\bar{\phi}(\bar{x}^1, \ldots, \bar{x}^n) = \phi(x^1, \ldots, x^n) \tag{3.7}$$

where $\bar{\phi}$ denotes the corresponding function expressed in terms of the coordinates $\bar{x}^j$ as defined by the orthonormal basis $\{\bar{\mathbf{e}}_j\}$. Under these circumstances $\phi(x^h)$ is called an *affine scalar*, or an *affine invariant*, and, in the present context, we shall regard such quantities as *affine tensors of rank zero*.

*Remark 2.* Because of (3.2) it is immediately evident that the inverse of (3.6) exists and can be written in the form

$$T_{h_1 h_2 \ldots h_r} = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_r=1}^{n} a_{j_1 h_1} a_{j_2 h_2} \cdots a_{j_r h_r} \bar{T}_{j_1 j_2 \ldots j_r}$$

$$(h_1, h_2, \ldots, h_r = 1, \ldots, n). \tag{3.8}$$

*Remark 3.* From (3.4) and (3.5) we infer the validity of the relations (2.18) and (2.19), irrespective of dimensionality, so that (3.6) and (3.8) can also be expressed as

$$\bar{T}_{j_1 j_2 \ldots j_r} = \sum_{h_1=1}^{n} \sum_{h_2=1}^{n} \cdots \sum_{h_r=1}^{n} \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \frac{\partial \bar{x}^{j_2}}{\partial x^{h_2}} \cdots \frac{\partial \bar{x}^{j_r}}{\partial x^{h_r}} T_{h_1 h_2 \ldots h_r}, \tag{3.9}$$

and

$$T_{h_1 h_2 \ldots h_r} = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_r=1}^{n} \frac{\partial x^{h_1}}{\partial \bar{x}^{j_1}} \frac{\partial x^{h_2}}{\partial \bar{x}^{j_2}} \cdots \frac{\partial x^{h_r}}{\partial \bar{x}^{j_r}} \bar{T}_{j_1 j_2 \ldots j_r}. \tag{3.10}$$

The following observation is of profound importance. An affine tensor is said to vanish, if, relative to a given orthonormal system, *all* its components are zero:

$$T_{h_1 h_2 \ldots h_r} = 0 \qquad (h_1, h_2, \ldots, h_r = 1, \ldots, n). \tag{3.11}$$

From the transformation law (3.6) it then follows that

$$\bar{T}_{j_1 j_2 \ldots j_r} = 0 \qquad (j_1, j_2, \ldots, j_r = 1, \ldots, n) \tag{3.12}$$

in any other rectangular coordinate system obtained from the first by means of the orthogonal transformation (3.4).

Thus *if an affine tensor vanishes in a given coordinate system, it vanishes in all other systems related to the first by an orthogonal transformation.*

Accordingly, whenever a geometrical or physical situation is described in a given rectangular coordinate system by means of a tensor equation such as (3.11), the same state of affairs is described in any other admissible coordinate system by the corresponding tensor equation (3.12); it follows that the validity of the assertion contained in (3.11) is independent of the choice of coordinates (provided that they are obtained from each other by an orthogonal transformation). This is precisely what one would expect of a geometrical theorem or of a physical law.

There are certain operations of an algebraic nature which one can perform with tensors. We deal here very briefly with operations of this kind; an alternative discussion is presented later under more general conditions.

Firstly, when the components of an affine tensor of given rank $r$ are each multiplied by the same nonzero number or scalar, the resulting quantities are once more the components of an affine tensor of the same rank $r$. This is immediately evident from the transformation law (3.6) when the right-hand side of the latter is multiplied by a nonvanishing scalar $\phi$, which, by virtue of (3.7), is tantamount to multiplication of its left-hand side by $\bar{\phi}$.

Secondly, when the respective components of two affine tensors *of the same rank* are added, the resulting sums are again the components of an affine tensor of the same rank. This follows directly from the fact that the left- and right-hand sides of the transformation law (3.6) are linear in the various components. For instance, let $T_{jkh}$ and $S_{jkh}$ denote the components of two tensors of rank 3, so that their transforms under (3.4) are respectively given by

$$\bar{T}_{jkh} = \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} a_{kp} a_{hq} T_{lpq}, \qquad \bar{S}_{jkh} = \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} a_{kp} a_{hq} S_{lpq}. \qquad (3.13)$$

Thus if we write

$$W_{lpq} = T_{lpq} + S_{lpq}, \qquad \bar{W}_{jkh} = \bar{T}_{jkh} + \bar{S}_{jkh}, \qquad (3.14)$$

in our respective coordinate systems, we find by means of (3.13) that

$$\bar{W}_{jkh} = \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} a_{kp} a_{hq} (T_{lpq} + S_{lpq}) = \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} a_{kp} a_{hq} W_{lpq}, \qquad (3.15)$$

which shows that the quantities $W_{lpq}$ are in fact the components of an affine tensor of rank 3.

From these two statements we infer that *the set of all affine tensors of rank $r$ constitutes a vector space over the real numbers.* The dimensionality of this vector space is equal to the number of distinct components of these tensors, namely, $n^r$. In particular, the set of all affine vectors forms an $n$-dimensional vector space.

From this it is immediately evident that *the addition of tensors of different rank cannot possibly give rise to a tensorial quantity.* For example, if one were to add one of the equations (3.13) to the transformation law of a rank 2 tensor, say,

$$\bar{V}_{jk} = \sum_{l=1}^{n} \sum_{p=1}^{n} a_{jl} a_{kp} V_{lp}, \qquad (3.16)$$

the result cannot be put into tensorial form. Indeed, while a combination such as $V_{jk} + U_{jk}$ makes sense in that the indices $j$, $k$ on $V$ and $U$ respectively are tacitly assumed to have the same values, this is not the case for a combination such as $V_{jk} + T_{jkh}$, because the index $h$ on $T$ has no counterpart on $V$, which therefore precludes an identification of this kind.

It is, however, permissible to multiply the components of tensors of arbitrary ranks, for this process gives rise to new tensors. For instance, if we were to multiply the left-hand side of (3.16) by the component $\bar{A}_h$ of an affine vector, this component being given by $\bar{A}_h = \sum_{q=1}^{n} a_{hq} A_q$, we would obtain

$$\bar{V}_{jk} \bar{A}_h = \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} a_{kp} a_{hq} V_{lp} A_q,$$

which shows that the quantities $V_{lp} A_q$ form the components of an affine tensor of rank 3.

It is easily verified that, in general, the quantities consisting of the products of the components of two tensors of rank $r$ and $s$ respectively again constitute the components of a tensor of rank $r + s$.

There is another operation, peculiar to tensor algebra, which is of considerable importance in certain manipulations. For any affine tensor of rank $r \geq 2$, say, $T_{j_1 j_2 \cdots j_p \cdots j_q \cdots j_r}$, one may select a pair of indices, say, $j_p$ and $j_q$, assign to these pairs of indices in turn the values $(1, 1), (2, 2), \ldots, (n, n)$, and form the sums of the $n$ components thus obtained. These quantities constitute the components of an affine tensor of rank $r - 2$. This is easily verified directly by means of (3.2) when the operation indicated above is carried out in respect of the transformation law (3.6). This process is generally referred to as the *contraction* of tensors, and we shall illustrate it for the case of the transformation law (3.13) of a tensor of rank 3.

Suppose, then, that we put $k = h$ on the left-hand side of (3.13), after which we sum over $h$:

$$\bar{T}_{j11} + \bar{T}_{j22} + \cdots + \bar{T}_{jnn} = \sum_{h=1}^{n} \bar{T}_{jhh} = \sum_{h=1}^{n} \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} a_{hp} a_{hq} T_{lpq}.$$

We then apply the orthogonality condition (3.2) to the sum $\sum_{h=1}^{n} a_{hp} a_{hq}$, thus obtaining

$$\sum_{h=1}^{n} \bar{T}_{jhh} = \sum_{l=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} a_{jl} \delta_{pq} T_{lpq} = \sum_{l=1}^{n} a_{jl}\left(\sum_{p=1}^{n} T_{lpp}\right),$$

which indicates that the quantities $\sum_{p=1}^{n} T_{lpp}$ form the components of an affine tensor of rank $1 = 3 - 2$.

It can be stated as a general rule that the *process of contraction applied once to a tensor of rank r gives rise to a new tensor of rank r − 2*.

In particular, if the process of contraction is applied to an affine tensor of rank 2, the result is an affine scalar. This establishes directly, for instance, the invariant nature of the trace $C_{11} + C_{22} + \cdots + C_{nn}$ of a matrix $(C_{hj})$, whose entries are the components of an affine tensor.

As a final example we note that, if $I_{jk}$ denotes the components of a tensor of rank 2, while $\omega_h$, $\omega_l$ are the components of a vector, the quantities $I_{jk}\omega_h\omega_l$ constitute the components of a tensor of rank 4. Contraction over the indices $(j, h)$ yields a rank 2 tensor, and subsequent contraction over $(l, k)$ gives rise to a scalar. Thus the invariant nature of the expression (1.1.13), namely,

$$\frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} I_{jk} \omega_j \omega_k \quad (n = 3),$$

is established effortlessly without calculation.

## 2.4  TRANSITION TO NONLINEAR COORDINATE TRANSFORMATIONS

It was stressed repeatedly that the tensor concept with which we have been concerned thus far is restricted in the sense that we are dealing with *affine* tensors which are defined relative to orthogonal transformations, the latter being *linear* by definition.

However, it is imperative that we should divest ourselves of this restriction. First, it is often necessary from a purely practical point of view to use curvilinear coordinates, the transition to which involves nonlinear coordinate transformations. Second, more often than not, one is concerned with tensor analysis on manifolds (such as curved surfaces) on which orthonormal coordinate systems simply cannot be defined; accordingly the theory developed thus far is entirely inadequate for situations of this kind. This observation, far more so than the first, provides the motivation for the type of generalization which we now attempt to achieve.

Let us endeavor to pinpoint those features of our earlier analysis which proved to be the most essential for the objectives thus far attained. One of these is the fact that the transformation law of tensors is always *linear in the*

*components of the tensors*: this allows us to define the processes of multiplication by scalars and addition of tensors of the same rank, which in turn permits us to regard all affine tensors of a given rank as elements of a vector space. Also, the success of the operation of contraction depends vitally on the relation (3.2) satisfied by the coefficients $a_{jh}$ which appear in the transformation law. It is evident, therefore, that if we can extend our theory such as to preserve these two features, a great deal of our basic analysis can be generalized as required.

We shall consider two coordinate systems in $E_n$, relative to which the coordinates of a point $P$ are denoted by $(x^1, \ldots, x^n)$ and $(\bar{x}^1, \ldots, \bar{x}^n)$, respectively. These two $n$-tuples of real numbers are related to each other by transformation equations of the type

$$\begin{cases} \bar{x}^1 = f^1(x^1, \ldots, x^n), \\ \cdots\cdots\cdots\cdots\cdots\cdots, \\ \bar{x}^n = f^n(x^1, \ldots, x^n), \end{cases}$$

in which the functions $f^1, \ldots, f^n$ are $n$ distinct functions of $n$ variables. In future we shall denote the $n$-tuple $(x^1, \ldots, x^n)$ by a single letter $x^h$, it being understood that all indices $j, h, k, \ldots$ assume the values 1 to $n$. Similarly we shall also write the above system of $n$ transformation equations in the form of a single relation, namely, as

$$\bar{x}^j = f^j(x^h) \qquad (j = 1, \ldots, n).$$

Furthermore, since the numerical value of each $\bar{x}^j$ is given by $f^j$, there is no need for two distinct symbols $\bar{x}$ and $f$, and accordingly, in order to economize on symbols, our transformation formulae henceforth are written in the form

$$\bar{x}^j = \bar{x}^j(x^h). \tag{4.1}$$

It is *not* assumed that the functions $\bar{x}^j(x^h)$ are necessarily linear in the coordinates $x^h$ [as is the case in (3.4), which is a special case of (4.1)]. At this stage, only two assumptions are made:

1.   The functions $\bar{x}^j(x^h)$ are of class $C^2$ in a region $G$ of $E_n$. [A function $g(x^h)$ is said to be of class $C^p$ in $G$ if it possesses continuous partial derivatives up to and including the $p$th order with respect to all its arguments $x^h$ in $G$.]

2.   The functional determinant

$$\frac{\partial(\bar{x}^1, \ldots, \bar{x}^n)}{\partial(x^1, \ldots, x^n)} = \begin{vmatrix} \dfrac{\partial \bar{x}^1}{\partial x^1} & \cdots & \dfrac{\partial \bar{x}^1}{\partial x^n} \\ \cdots\cdots\cdots\cdots \\ \dfrac{\partial \bar{x}^n}{\partial x^1} & \cdots & \dfrac{\partial \bar{x}^n}{\partial x^n} \end{vmatrix} \tag{4.2}$$

is nonvanishing on a subregion of $G$; this implies (see, e.g., Apostol [1]) that there exists a region $R \subseteq G$ in which the transformation (4.1) possesses an inverse, expressed in the form

$$x^h = x^h(\bar{x}^j), \tag{4.3}$$

it again being understood that this represents a system of $n$ equations, of which the right-hand sides are functions of the $n$ variables $\bar{x}^j$. Indeed, by virtue of the inverse function theorem it follows from our assumptions that these functions are also of class $C^2$.

*Henceforth our considerations will be restricted to the region $R$ in which the inverse (4.3) exists*: this is to be understood throughout without specific mention.

For future reference we note the following elementary fact. Let us substitute the functions $\bar{x}^j(x^h)$ in (4.3), which yields a system of identities in the variables $x^k$, namely,

$$x^h = x^h(\bar{x}^j(x^k)),$$

which, written out more fully, is of the form

$$x^1 = x^1(\bar{x}^1(x^1, \ldots, x^n), \ldots, \bar{x}^n(x^1, \ldots, x^n)),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots,$$
$$x^n = x^n(\bar{x}^1(x^1, \ldots, x^n), \ldots, \bar{x}^n(x^1, \ldots, x^n)).$$

By means of the chain rule we may differentiate each of these identities with respect to $x^k$. For the $h$th identity we thus obtain

$$\frac{\partial x^h}{\partial x^k} = \frac{\partial x^h}{\partial \bar{x}^1}\frac{\partial \bar{x}^1}{\partial x^k} + \cdots + \frac{\partial x^h}{\partial \bar{x}^n}\frac{\partial \bar{x}^n}{\partial x^k} = \sum_{j=1}^n \frac{\partial x^h}{\partial \bar{x}^j}\frac{\partial \bar{x}^j}{\partial x^k}.$$

But the variables $x^1, \ldots, x^n$ are independent, so that $\partial x^h/\partial x^k = \delta_{hk}$, and accordingly one has

$$\sum_{j=1}^n \frac{\partial x^h}{\partial \bar{x}^j}\frac{\partial \bar{x}^j}{\partial x^k} = \delta_{hk}. \tag{4.4}$$

Similarly

$$\sum_{h=1}^n \frac{\partial \bar{x}^j}{\partial x^h}\frac{\partial x^h}{\partial \bar{x}^l} = \delta_{jl}. \tag{4.5}$$

After these preparations we now recall that, in the case of a linear transformation, the transformation law (3.6) of an affine tensor can be expressed in the form (3.9), which is possible because of the identifications (2.18) and (2.19)

of the coefficients $a_{jk}$ with the partial derivatives of the relevant *coordinate* transformation, namely, $a_{jk} = \partial \bar{x}^j/\partial x^k$ and $a_{jh} = \partial x^h/\partial \bar{x}^j$. Obviously this preserves the linear character of the tensor transformation law as far as the tensor components are concerned; thus since the partial derivatives $\partial \bar{x}^j/\partial x^k$ are at our disposal in any case, this suggests that we should formulate the transformation law for tensors in respect of the general coordinate transformation (4.1) in terms of these derivatives exactly as in (3.9). Furthermore, the counterpart of the orthogonality condition (3.2) assumes the form (4.4) or (4.5) after the replacement of the coefficients $a_{jh}$ by the $\partial \bar{x}^j/\partial x^h$ has been made, and, as we have seen above, *in this form the crucial condition (3.2) is always satisfied*.

From a purely analytical point of view this suggestion seems to be a satisfactory one, at least at first sight. It implies that we should, for instance, define a vector by the requirement that its components $A_h$ should transform according to the relation

$$\bar{A}_j = \sum_{h=1}^n \frac{\partial \bar{x}^j}{\partial x^h} A_h \tag{4.6}$$

relative to coordinate transformation (4.1). However, before proceeding further, we should verify that this program is in accordance with our geometrical approach of Section 1.2 to the concept of the components of a vector in curvilinear coordinates.

We shall therefore consider a special case in $E_3$, namely, our *purely geometrical* construction of the components of a vector $\mathbf{A}$ in spherical polar coordinates. In this context we shall denote the latter as follows:

$$\bar{x}^1 = \rho, \qquad \bar{x}^2 = \theta, \qquad \bar{x}^3 = \phi, \tag{4.7}$$

and the transformation (1.2.1) from these curvilinear coordinates to the appropriate rectangular coordinates is given by

$$\left.\begin{array}{l} x^1 = \bar{x}^1 \sin \bar{x}^2 \cos \bar{x}^3, \\ x^2 = \bar{x}^1 \sin \bar{x}^2 \sin \bar{x}^3, \\ x^3 = \bar{x}^1 \cos \bar{x}^2, \end{array}\right\} \tag{4.8}$$

this being a special case of (4.3). The matrix of the derivatives $\partial x^h/\partial \bar{x}^j$ of this transformation is

$$\left(\frac{\partial x^h}{\partial \bar{x}^j}\right) = \begin{pmatrix} \sin \bar{x}^2 \cos \bar{x}^3 & \bar{x}^1 \cos \bar{x}^2 \cos \bar{x}^3 & -\bar{x}^1 \sin \bar{x}^2 \sin \bar{x}^3 \\ \sin \bar{x}^2 \sin \bar{x}^3 & \bar{x}^1 \cos \bar{x}^2 \sin \bar{x}^3 & \bar{x}^1 \sin \bar{x}^2 \cos \bar{x}^3 \\ \cos \bar{x}^2 & -\bar{x}^1 \sin \bar{x}^2 & 0 \end{pmatrix}. \tag{4.9}$$

A simple calculation yields the inverse:

$$\left(\frac{\partial \bar{x}^j}{\partial x^h}\right) = \begin{pmatrix} \sin \bar{x}^2 \cos \bar{x}^3 & \sin \bar{x}^2 \sin \bar{x}^3 & \cos \bar{x}^2 \\ \dfrac{\cos \bar{x}^2 \cos \bar{x}^3}{\bar{x}^1} & \dfrac{\cos \bar{x}^2 \sin \bar{x}^3}{\bar{x}^1} & -\dfrac{\sin \bar{x}^2}{\bar{x}^1} \\ -\dfrac{\sin \bar{x}^3}{\bar{x}^1 \sin \bar{x}^2} & \dfrac{\cos \bar{x}^3}{\bar{x}^1 \sin \bar{x}^2} & 0 \end{pmatrix}. \tag{4.10}$$

Thus, substituting from (4.10) in (4.6) we infer that, according to this prescription, the components $\bar{A}_j$ of the vector $\mathbf{A}$ are given in spherical polar coordinates by

$$\bar{A}_1 = \sin \bar{x}^2 \cos \bar{x}^3 A_1 + \sin \bar{x}^2 \sin \bar{x}^3 A_2 + \cos \bar{x}^2 A_3,$$

$$\bar{A}_2 = \frac{\cos \bar{x}^2 \cos \bar{x}^3}{\bar{x}^1} A_1 + \frac{\cos \bar{x}^2 \sin \bar{x}^3}{\bar{x}^1} A_2 - \frac{\sin \bar{x}^2}{\bar{x}^1} A_3, \tag{4.11}$$

$$\bar{A}_3 = -\frac{\sin \bar{x}^3}{\bar{x}^1 \sin \bar{x}^2} A_1 + \frac{\cos \bar{x}^3}{\bar{x}^1 \sin \bar{x}^2} A_2.$$

This is seen to coincide with our previous expressions (1.2.7) for the components of $\mathbf{A}$ in spherical polar coordinates when we revert to the original notation by means of (4.7). Accordingly this conclusion seems to confirm, at least from a geometrical point of view, that the suggested definition based on (4.6) is a satisfactory one.

Nevertheless, a little reflection shows that our approach requires some additional refinements. For instance, if we consider a set of components $B_k$, satisfying the transformation law (4.6), namely,

$$\bar{B}_l = \sum_{k=1}^{n} \frac{\partial \bar{x}^l}{\partial x^k} B_k, \tag{4.12}$$

the products of (4.6) and (4.12) yield

$$\bar{A}_j \bar{B}_l = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^l}{\partial x^k} A_h B_k, \tag{4.13}$$

which exemplifies what would seem to be an acceptable transformation law for a rank 2 tensor. However, if we contract over the indices $j, l$ (in an attempt to obtain an inner product), we find that

$$\sum_{j=1}^{n} \bar{A}_j \bar{B}_j = \sum_{j=1}^{n} \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^j}{\partial x^k} A_h B_k, \tag{4.14}$$

in which the expression on the right-hand side cannot, in general, be simplified

by an application of (4.4) or (4.5). Thus the process of contraction, as carried out in this context, does not give rise to a corresponding scalar. [The fact that this difficulty does not arise in the case of linear orthogonal coordinate transformations is due to the fact that, according to (2.18) and (2.19), we have

$$\frac{\partial \bar{x}^j}{\partial x^h} = a_{jh} = \frac{\partial x^h}{\partial \bar{x}^j} \tag{4.15}$$

for such transformations; if it were possible to substitute (4.15) in (4.14) the identity (4.4) could be readily applied.]

Clearly our program still lacks some vital ingredient, for it is essential that somehow one should be able to construct scalars by means of suitable combinations of tensors. Thus we approach this problem from a slightly different angle by first considering scalars and certain vectors associated with scalars.

A function $\phi(x^h)$ of the coordinates $x^h$ is said to be a *scalar* or *invariant* under the transformation (4.1) if its transform $\bar{\phi}(\bar{x}^j)$ possesses the same numerical value; that is, if

$$\bar{\phi}(\bar{x}^j) = \phi(x^h). \tag{4.16}$$

Here it is to be clearly understood that the arguments $\bar{x}^j$ are related to the arguments $x^h$ according to (4.1): both sets of coordinates refer to the *same* point $P$ of $E_n$.

With any differentiable scalar function $\phi$ one usually associates the so-called *gradient vector*: the $n$ components of the latter are defined to be the $n$ partial derivatives $\partial \phi / \partial x^h$. The question that immediately arises concerns the transformation properties of these derivatives, for, if they are to constitute the components of a vector in the sense of the proposed transformation law (4.6), the latter should be satisfied by $\partial \phi / \partial x^h$. In order to test this criterion, we differentiate (4.16) partially with respect to $\bar{x}^j$, the chain rule being applied to the right-hand side, which gives

$$\frac{\partial \bar{\phi}}{\partial \bar{x}^j} = \sum_{h=1}^{n} \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \phi}{\partial x^h}. \tag{4.17}$$

This would be in accordance with (4.6) only if one could identify $\partial x^h / \partial \bar{x}^j$ with $\partial \bar{x}^j / \partial x^h$, which is not generally possible. Thus (4.17) exemplifies a transformation law of the type

$$\bar{C}_j = \sum_{h=1}^{n} \frac{\partial x^h}{\partial \bar{x}^j} C_h, \tag{4.18}$$

which is distinct from (4.6). However, the example (4.17) is obviously a significant one which cannot be ignored. Furthermore, if we combine two sets of

components, the first of which satisfy (4.6), the second (4.18), we obtain quantities of the type

$$\bar{A}_j \bar{C}_l = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^l} A_h C_k, \tag{4.19}$$

and if the process of contraction over $j$ and $l$ is carried out in this instance we obtain, using (4.4) on the right-hand side,

$$\sum_{j=1}^{n} \bar{A}_j \bar{C}_j = \sum_{j=1}^{n} \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^j} A_h C_k = \sum_{h=1}^{n} \sum_{k=1}^{n} \delta_{hk} A_h C_k = \sum_{h=1}^{n} A_h C_h. \tag{4.20}$$

In contrast to (4.14) this result is indeed very satisfactory, for it indicates that we have succeeded in constructing an invariant resembling the inner product of two vectors.

This state of affairs indicates quite clearly that we must consider two types of transformation laws, namely, (4.6) and (4.18), and accordingly *we must define two distinct kinds of vectors.* [This distinction does not arise in the theory of linear orthogonal transformations since in the case of the latter the condition (4.15) is satisfied which deletes the distinction between (4.6) and (4.18).]

As regards nomenclature, we shall distinguish between these two kinds of vectors by calling the former *contravariant* and the latter *covariant*. This distinction must also be reflected in our notation: *henceforth contravariant vectors are endowed with superscripts, and covariant vectors are distinguished by subscripts.* This notation is used consistently; it forms an essential aspect of tensor calculus.

The results of these considerations may now be crystallized in the following two definitions.

**DEFINITION**

*A set of $n$ quantities $(A^1, \ldots, A^n)$ is said to constitute the components of a contravariant vector at a point $P$ with coordinates $(x^1, \ldots, x^n)$ if, under the transformation (4.1), these quantities transform according to the relations*

$$\bar{A}^j = \sum_{h=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} A^h, \tag{4.21}$$

*in which the coefficients $\partial \bar{x}^j / \partial x^h$ are to be evaluated at $P$.*

**DEFINITION**

*A set of $n$ quantities $(C_1, \ldots, C_n)$ is said to constitute the components of a covariant vector at a point $P$ with coordinates $(x^1, \ldots, x^n)$ if, under the trans-*

*formation (4.1), these quantities transform according to the relations*

$$\bar{C}_j = \sum_{h=1}^{n} \frac{\partial x^h}{\partial \bar{x}^j} C_h, \tag{4.22}$$

*in which the coefficients $\partial x^h / \partial \bar{x}^j$ are to be evaluated at $P$.*

*Remark 1.* For future reference we note that co- and contravariant vectors may also be called type (0, 1) and type (1, 0) tensors, respectively.

*Remark 2.* These definitions contain as a special case the definitions of an *affine* vector in $E_3$: for, when the transformation (4.1) happens to be an orthogonal one with $n = 3$, both (4.21) and (4.22) reduce to (2.11).

*Remark 3.* Several of the equations above should now be rewritten with various subscripts replaced by superscripts; in particular, the identity (4.20) is to be expressed in the form

$$\sum_{j=1}^{n} \bar{A}^j \bar{C}_j = \sum_{h=1}^{n} A^h C_h. \tag{4.23}$$

*Remark 4.* The coefficients $\partial \bar{x}^j / \partial x^h$ which appear in the transformation law (4.21) are, in general, functions of the variables $x^1, \ldots, x^n$, which are the positional coordinates of the point of application of the contravariant vector $\mathbf{A}$. This implies that *one can add two contravariant vectors $\mathbf{A}$ and $\mathbf{B}$ if and only if they are located at the same point.* [For instance, if the points of application of the vectors are distinct, with coordinates $x^k = x^k_{(1)}$ and $x^k = x^k_{(2)}$, respectively, the corresponding transformation equations read

$$\bar{A}^j = \sum_{h=1}^{n} \frac{\partial \bar{x}^j(x^k_{(1)})}{\partial x^h} A^h \quad \text{and} \quad \bar{B}^j = \sum_{h=1}^{n} \frac{\partial \bar{x}^j(x^k_{(2)})}{\partial x^h} B^h;$$

since the values of the coefficients of $A^h$ and $B^h$ on the right-hand sides do not in general coincide, the sum of the right-hand sides is not linear in $A^h + B^h$.] It is, however, possible to multiply both sides of (4.21) by a scalar function $\phi(x^h)$ since the latter satisfies (4.16), it being assumed that the arguments $x^1, \ldots, x^n$ in $\phi(x^h)$ also refer to the point of application of $\mathbf{A}$. *Thus the set of all contravariant vectors at a point $P$ of $E_n$ constitutes a vector space of dimension $n$.* Similar observations apply, of course, to covariant vectors. Moreover, it is evident that *the addition of co- and contravariant vectors cannot give rise to a vectorial quantity.*

As in the case of affine vectors one can form the products of the respective components of various vectors (provided that these vectors are located at the same point), thus obtaining the components of quantities which will be recognized as tensors of rank $r > 1$. Again, a clear distinction must be made

between co- and contravariant properties of the entities thus constructed. For instance, given two contravariant vectors with components $A^h$ and $B^k$ located at the same point $P$ of $E_n$, it follows from (4.21) that the transforms of the products of their components are given by

$$\bar{A}^j \bar{B}^l = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^l}{\partial x^k} A^h B^k,$$

which exemplifies the *transformation law of a tensor of type* (2, 0):

$$\bar{T}^{jl} = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^l}{\partial x^k} T^{hk}. \tag{4.24}$$

Similarly (4.22) suggests the *transformation law of a tensor of type* (0, 2):

$$\bar{S}_{jl} = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial x^k}{\partial \bar{x}^l} S_{hk}. \tag{4.25}$$

However, it is also feasible to define the so-called "mixed" tensors, possessing both co- and contravariant properties. For instance, if we form the products of the components $A^h$ of a contravariant vector with the components $C_k$ of a covariant vector, it follows directly from (4.21) and (4.22) that the quantities thus obtained transform according to

$$\bar{A}^j \bar{C}_l = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^l} A^h C_k. \tag{4.26}$$

This exemplifies the *transformation law of a type* (1, 1) *tensor*, namely,

$$\bar{V}_l^j = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^l} V_k^h. \tag{4.27}$$

Remarks 2 and 4 made above in respect of the definitions of co- and contravariant vectors apply more or less verbatim to the tensor components which appear in (4.24), (4.25), and (4.27).

Furthermore, by analogy with the theory of affine tensors it is now obvious how one can extend these transformation equations to tensors of arbitrary type. We shall defer these formal definitions to the next chapter.

In the sequel we encounter numerous examples of the various tensors discussed here. However, already at this stage we should draw attention to a mixed tensor of special importance. Bearing in mind that, by definition, the Kronecker delta assumes the numerical values 0 and 1 irrespective of the choice of coordinates, and writing (4.5) in the form

$$\bar{\delta}_{jl} = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^l} \delta_{hk},$$

we see by comparison with (4.27) that *the Kronecker delta is, in fact, a type* (1, 1) *tensor*. Accordingly we henceforth denote its components by $\delta_k^h$, and the notationally correct version of its transformation law is

$$\bar{\delta}_l^j = \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^l} \delta_k^h, \tag{4.28}$$

while (4.4) and (4.5) should be written in the forms

$$\sum_{j=1}^{n} \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \bar{x}^j}{\partial x^k} = \delta_k^h, \tag{4.29}$$

and

$$\sum_{h=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^l} = \delta_l^j. \tag{4.30}$$

## 2.5  DIGRESSION: PARALLEL VECTOR FIELDS IN $E_n$ REFERRED TO CURVILINEAR COORDINATES

In an $n$-dimensional Euclidean space $E_n$ let us consider two coordinate systems in which the coordinates of a point $P$ are denoted by $\bar{x}^j$ and $x^h$, respectively. It is assumed that the first of these is a rectangular coordinate system, whereas the second is an arbitrary curvilinear system, the corresponding transformation being of the form (4.1), for which the functions $\bar{x}^j(x^h)$ are now supposed to be of class $C^2$.

It is possible to construct a special vector field in $E_n$ which is such that there exists a unique vector at each point of $E_n$, these vectors being parallel and of the same length. The components $\bar{X}^j$ of this field relative to the rectangular system are obviously constant; and thus, for any displacement $d\bar{x}^k$ the conditions

$$d\bar{X}^j = \sum_{k=1}^{n} \frac{\partial \bar{X}^j}{\partial \bar{x}^k} d\bar{x}^k = 0 \tag{5.1}$$

must be satisfied.

The components $X^h$ of the same vector field relative to the curvilinear system are related to the $\bar{X}^j$ according to

$$\bar{X}^j = \sum_{h=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} X^h, \tag{5.2}$$

so that, corresponding to any displacement $dx^k$, we have

$$d\bar{X}^j = \sum_{h=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} dX^h + \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} X^h \, dx^k. \tag{5.3}$$

From this it is evident that the differentials of a vector field are not, in general, components of a tensor: furthermore, the condition (5.1) for a parallel vector field in $E_n$ does *not* imply a corresponding relation of the type $dX^h = 0$. Thus relative to our curvilinear coordinate system, a field of parallel vectors of constant length is *not* characterized by the requirement that the components of this field be constant.

Thus the question arises as to how such vector fields are to be described in arbitrary curvilinear coordinates, and we now give a complete analysis of this problem. To this end we recall that, relative to our rectangular coordinate system, the length $|\mathbf{X}|$ of any vector $\mathbf{X}$ of our field in $E_n$ is given by

$$|\mathbf{X}|^2 = \sum_{j=1}^{n} \bar{X}^j \bar{X}^j. \tag{5.4}$$

By means of (5.2) this can be written in terms of our curvilinear coordinate system as follows:

$$|\mathbf{X}|^2 = \sum_{h=1}^{n} \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^j}{\partial x^k} X^h X^k. \tag{5.5}$$

This suggests that we introduce new quantities $g_{hk}$ defined by

$$g_{hk} = \sum_{j=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^j}{\partial x^k}, \tag{5.6}$$

so that (5.5) can be written in the form

$$|\mathbf{X}|^2 = \sum_{h=1}^{n} \sum_{k=1}^{n} g_{hk} X^h X^k. \tag{5.7}$$

Again this demonstrates the fact that the length of a contravariant vector at $P$ is a quadratic form in the components of that vector, the coefficients $g_{hk}$ of this quadratic form as defined by (5.6) being evaluated at $P$. [In fact, the relation (5.7) is the generalization to $n$ dimensions of the identity (1.2.16).]

For future reference we note that the definition (5.6) implies that $\det(g_{hk}) \neq 0$ by virtue of our assumption that the functional determinant (4.2) does not vanish. Thus the symmetric matrix $(g_{hk})$ possesses an inverse, to be denoted by $(g^{hk})$, so that

$$\sum_{l=1}^{n} g^{jl} g_{hl} = \delta_h^j. \tag{5.8}$$

We shall now endeavor to eliminate the second derivatives which appear on the right-hand side of (5.3) by means of the $g_{hk}$ in the following manner. First, multiplying (5.3) by $\partial \bar{x}^j/\partial x^l$ and summing over $j$, at the same time noting

(5.6), we obtain

$$\sum_{j=1}^{n} \frac{\partial \bar{x}^j}{\partial x^l} d\bar{X}^j = \sum_{h=1}^{n} g_{hl} dX^h + \sum_{j=1}^{n} \sum_{h=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} X^h dx^k. \tag{5.9}$$

Second, we differentiate (5.6) partially with respect to $x^l$, which yields

$$\frac{\partial g_{hk}}{\partial x^l} = \sum_{j=1}^{n} \left( \frac{\partial^2 \bar{x}^j}{\partial x^l \partial x^h} \frac{\partial \bar{x}^j}{\partial x^k} + \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 \bar{x}^j}{\partial x^l \partial x^k} \right). \tag{5.10}$$

The second derivatives thus obtained are not yet expressed in a form most suitable for our purposes; accordingly we consider two cyclic permutations of the indices $h$, $k$, $l$ in (5.10), corresponding to which the counterparts of (5.10) are

$$\frac{\partial g_{kl}}{\partial x^h} = \sum_{j=1}^{n} \left( \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} + \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^l} \right), \tag{5.11}$$

and

$$\frac{\partial g_{lh}}{\partial x^k} = \sum_{j=1}^{n} \left( \frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^l} \frac{\partial \bar{x}^j}{\partial x^h} + \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^h} \right). \tag{5.12}$$

In order to reduce (5.9) to an acceptable form we seek the expression

$$\sum_{j=1}^{n} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^j}{\partial x^l},$$

which appears in both (5.11) and (5.12). These relations are added, and it is found that the unwanted terms may be removed by subtracting (5.10) from the sum thus obtained. In this way we find that

$$\left( \frac{\partial g_{kl}}{\partial x^h} + \frac{\partial g_{lh}}{\partial x^k} - \frac{\partial g_{hk}}{\partial x^l} \right) = 2 \sum_{j=1}^{n} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^j}{\partial x^l}. \tag{5.13}$$

The following notation is therefore introduced (in the first instance merely for the sake of brevity):

$$[hk, l] = \frac{1}{2} \left( \frac{\partial g_{kl}}{\partial x^h} + \frac{\partial g_{lh}}{\partial x^k} - \frac{\partial g_{hk}}{\partial x^l} \right), \tag{5.14}$$

these three-index quantities being called the *Christoffel symbols of the first kind of our curvilinear coordinate system*. Thus (5.13) can be expressed as

$$\sum_{j=1}^{n} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} = [hk, l], \tag{5.15}$$

and (5.9) becomes

$$\sum_{j=1}^{n} \frac{\partial \bar{x}^j}{\partial x^l} d\bar{X}^j = \sum_{h=1}^{n} g_{hl} dX^h + \sum_{h=1}^{n} \sum_{k=1}^{n} [hk, l] X^h dx^k. \tag{5.16}$$

Now, if it is assumed that the vector field $\bar{X}^j$ is parallel and of constant length in $E_n$, we may apply the criterion (5.1) to the left-hand side of (5.16), which yields the condition

$$\sum_{h=1}^{n} g_{hl} dX^h + \sum_{h=1}^{n} \sum_{k=1}^{n} [hk, l] X^h dx^k = 0 \tag{5.17}$$

for such a vector field in curvilinear coordinates. *This condition is expressed entirely in terms of quantities which pertain to this particular coordinate system.*

In order to write (5.17) in a slightly better form, we multiply it by $g^{jl}$, the inverse of (5.6), and sum over $l$, at the same time noting (5.8). This gives

$$dX^j + \sum_{l=1}^{n} \sum_{h=1}^{n} \sum_{k=1}^{n} g^{jl}[hk, l] X^h dx^k = 0, \tag{5.18}$$

which suggests the notation

$$\{_h{}^j{}_k\} = \sum_{l=1}^{n} g^{jl}[hk, l], \tag{5.19}$$

for the so-called *Christoffel symbols of the second kind of our curvilinear coordinate system.* Accordingly we can write (5.18) in the form

$$dX^j + \sum_{h=1}^{n} \sum_{k=1}^{n} \{_h{}^j{}_k\} X^h dx^k = 0, \tag{5.20}$$

*which is the condition in curvilinear coordinates for a field of parallel vectors of constant length.*

For many practical purposes it is advisable to write (5.20) as a system of partial differential equations. For any displacement $dx^k$ we have

$$dX^j = \sum_{k=1}^{n} \frac{\partial X^j}{\partial x^k} dx^k, \tag{5.21}$$

which may be substituted in (5.20) to yield

$$\sum_{k=1}^{n} \left( \frac{\partial X^j}{\partial x^k} + \sum_{h=1}^{n} \{_h{}^j{}_k\} X^h \right) dx^k = 0,$$

and, since this must hold for entirely arbitrary displacements $dx^k$, it follows

that *our condition for the field of parallel vectors of constant length is equivalent to*

$$\frac{\partial X^j}{\partial x^k} + \sum_{h=1}^{n} \{_h{}^j{}_k\} X^h = 0, \tag{5.22}$$

*which is a system of $n^2$ partial differential equations of the first order to be satisfied by the components of the vector field.*

*Remark 1.* This condition is expressed entirely in terms of quantities defined relative to the curvilinear coordinate system. It holds in *any* curvilinear coordinate system; accordingly one would infer that the left-hand side of (5.22) is tensorial. (This will be verified by direct calculation at a later stage under more general circumstances.)

*Remark 2.* In the original rectangular coordinate system the counterparts $\bar{g}_{hk}$ of the quantities $g_{hk}$ are given by $\bar{g}_{hk} = \delta_{hk} = $ const., as is immediately evident by comparison of (5.4) with (5.7). *Thus the Christoffel symbols (5.14) and (5.19) vanish identically in any rectangular coordinate system in $E_n$,* and accordingly for such systems the condition (5.20) reduces formally to (5.1).

*Remark 3.* As a simple application of the above analysis let us consider the equation of a straight line in $E_n$. Relative to the rectangular coordinate system a straight line $\bar{x}^j = \bar{x}^j(s)$, where $s$ denotes the arc length, is characterized by the condition

$$\frac{d\bar{x}'^j}{ds} = 0, \tag{5.23}$$

in which we have written

$$\bar{x}'^j = \frac{d\bar{x}^j}{ds}. \tag{5.24}$$

In (5.16), which holds for any vector $\bar{X}^j$, let us put $\bar{X}^j = \bar{x}'^j$, and divide by $ds$, observing (5.23) at the same time. This gives

$$\sum_{h=1}^{n} \left( g_{hl} \frac{dx'^h}{ds} + \sum_{k=1}^{n} [hk, l] x'^h \frac{dx^k}{ds} \right) = 0,$$

or, proceeding as before in the transition from (5.17) to (5.20),

$$\frac{d^2 x^j}{ds^2} + \sum_{h=1}^{n} \sum_{k=1}^{n} \{_h{}^j{}_k\} \frac{dx^h}{ds} \frac{dx^k}{ds} = 0. \tag{5.25}$$

*This system of $n$ second-order ordinary differential equations characterizes the straight lines of $E_n$ relative to any curvilinear coordinate system.*

## PROBLEMS

2.1   In $E_3$ a rotation through an angle $\alpha$ about the $x^3$-axis is characterized by the matrix

$$(a_{ij}) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Show that this is a proper orthogonal transformation.

A reflection in the plane $x^2 \cos\beta = x^1 \sin\beta$ is determined by

$$(b_{ij}) = \begin{pmatrix} \cos 2\beta & \sin 2\beta & 0 \\ \sin 2\beta & -\cos 2\beta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Show that this is an orthogonal transformation which is not proper.

2.2   The components $A_i$ of an affine vector in $E_3$ have the values $(\frac{1}{2}, 0, 0)$ in the $x^k$ coordinate system. If

$$(a_{jk}) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is the matrix of the (orthogonal) transformation from $x^i$ to $\bar{x}^j$, compute $\bar{A}_i$.

2.3   Prove that the set of all proper orthogonal transformations in $E_n$ forms a group, but that the set of all orthogonal transformations for which the associated determinant is $-1$ does not form a group.

2.4   In $E_n$ the quantity $\phi = \phi(x^h)$ is an affine scalar. Prove that $\partial^2\phi/(\partial x^i \partial x^j)$ are the components of an affine tensor of rank 2.

2.5   In $E_n$, $A_h$, $T_{ij}$, and $U_{lkm}$ are the components of affine tensors of rank 1, 2, and 3, respectively. If

$$A_h T_{ij} = U_{ihj}$$

in one rectangular coordinate system, establish that

$$\bar{A}_h \bar{T}_{ij} = \bar{U}_{ihj}$$

in any other rectangular coordinate system.

2.6   In $E_n$ prove that $\delta_{ij}$ are the components of an affine tensor of rank 2.

*2.7   In $E_n$ the quantities $B_{ij}$ are the components of an affine tensor of rank 2. Construct two affine tensors each of rank 4, with components $C_{ijkl}$ and $D_{ijkl}$ for which

$$\sum_{k=1}^{n} \sum_{l=1}^{n} C_{ijkl} B_{kl} = B_{ij} + B_{ji}$$

$$\sum_{k=1}^{n} \sum_{l=1}^{n} D_{ijkl} B_{kl} = B_{ij} - B_{ji}$$

are identities.

2.8   Show that, in $E_3$, the quantities $\varepsilon_{ijk}$ defined in Problem 1.1 do not constitute the components of an affine tensor of rank 3 unless the transformation is proper, whereas the quantities $\varepsilon_{ijk}\varepsilon_{hlm}$ are the components of an affine tensor of rank 6. Establish that

$$\sum_{j=1}^{3} \varepsilon_{ijj} = 0, \qquad \sum_{k=1}^{3} \varepsilon_{ijk}\varepsilon_{hlk} = \delta_{ih}\delta_{jl} - \delta_{il}\delta_{hj}$$

and

$$\sum_{k=1}^{3} \sum_{j=1}^{3} \varepsilon_{ijk}\varepsilon_{hjk} = 2\,\delta_{ih}.$$

2.9   In $E_3$ the components of the vector product $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ of two nonparallel vectors are given by

$$C_1 = (A_2 B_3 - A_3 B_2), \qquad C_2 = (A_3 B_1 - A_1 B_3), \qquad C_3 = (A_1 B_2 - A_2 B_1).$$

The left-hand sides, regarded as components of an affine vector transform according to (2.11), which is *linear* in the coefficients $a_{jh}$ of the orthogonal transformation. On the other hand, the right-hand sides, regarded as components of an affine tensor, transform according to (2.25), which is *quadratic* in the $a_{jh}$. Show how this apparent contradiction can be resolved provided that the transformation is proper.

2.10   The orthonormal basis $\{e_j\}$ in $E_3$ is "right-handed" in the sense that $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$, and $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$.

If $\{f_j\}$ represents an orthonormal basis obtained from $\{e_j\}$ according to (1.2), show that

$$\mathbf{f}_1 \times \mathbf{f}_2 = \sum_{h=1}^{3} A_{3h}\mathbf{e}_h,$$

where $A_{jh}$ denotes the cofactors in the determinant $a = \det(a_{jh})$. Hence deduce that

$$\mathbf{f}_1 \times \mathbf{f}_2 = +\mathbf{f}_3 \quad \text{or} \quad -\mathbf{f}_3$$

according as $a = +1$ or $-1$. (This result shows that proper orthogonal transformations preserve the "right-handedness" or orientation of orthonormal systems.)

2.11   If $A_i$ and $B_j$ are the components of two affine vectors in $E_n$ the angle $\theta$ between them is defined according to

$$\cos\theta = \frac{\sum_{i=1}^{n} A_i B_i}{\left(\sum_{j=1}^{n} A_j A_j \sum_{k=1}^{n} B_k B_k\right)^{1/2}}.$$

Prove that this quantity is an affine invariant.

2.12   In $E_n$, if the quantities $A_i = A_i(x^h)$ are the components of an affine vector show that $\partial A_i/\partial x^j$ are the components of an affine tensor of rank 2.

2.13   In $E_n$, the quantities $A_i$ are the components of an affine vector for which

$$\sum_{i=1}^{n} A_i A_i = 1.$$

If $P_{ij} = A_i A_j$ show that

$$\sum_{i=1}^{n} P_{ij} P_{ik} = P_{jk}.$$

If, for $n = 3$    (a) $P_{ij} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$$(b) \ P_{ij} = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix},$$

find $A_i$, if it exists.

*2.14 In $E_n$, $b_{ij}$ and $c_{ij}$ are the components of affine tensors of rank 2. If $c_{ij} = -c_{ji}$ and

$$\sum_{j=1}^{n} (c_{ij} b_{jk} + c_{kj} b_{ij}) = 0$$

for all $c_{ij}$, show that for $n > 2$

$$b_{ij} = \lambda \delta_{ij}.$$

Determine $\lambda$ and deduce that it is an affine scalar.

2.15 In $E_n$ the quantity $\phi = \phi(x^h)$ is a scalar under (4.1). Prove that $\partial^2 \phi / \partial x^i \, \partial x^j$ are not the components of a tensor of either type $(2, 0)$, $(1, 1)$, or $(0, 2)$. (Compare with Problem 2.4.)

2.16 Let $\bar{x}^j$, $\tilde{x}^j$ denote the coordinates of an arbitrary point $P$ of $E_n$ referred to two distinct rectangular coordinate systems. An arbitrary curvilinear system in $E_n$ is related to the two rectangular systems according to

$$\bar{x}^j = \bar{x}^j(x^h), \qquad \tilde{x}^j = \tilde{x}^j(x^h).$$

Show that

$$\sum_{j=1}^{n} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^j}{\partial x^k} = \sum_{j=1}^{n} \frac{\partial \tilde{x}^j}{\partial x^h} \frac{\partial \tilde{x}^j}{\partial x^k}.$$

Deduce that the quantities (5.6) defined for the curvilinear coordinate system are independent of the choice of the rectangular coordinate system.

2.17 Let $x^h$, $\bar{x}^h$ denote the coordinates of an arbitrary point $P$ of $E_n$ referred to two arbitrary curvilinear coordinate systems. These coordinates are related to each other by the transformation equations $\bar{x}^h = \bar{x}^h(x^k)$. Let the quantities (5.6) defined for each of these systems be denoted by $g_{hk}$, $\bar{g}_{hk}$, respectively. Show that

$$\bar{g}_{hk} = \sum_{j=1}^{n} \sum_{l=1}^{n} \frac{\partial x^j}{\partial \bar{x}^h} \frac{\partial x^l}{\partial \bar{x}^k} g_{jl},$$

which indicates that the $g_{jl}$ are components of a type $(0, 2)$ tensor. Deduce the transformation law satisfied by $g = \det(g_{jl})$. Is the latter a tensor?

2.18 Find explicit expressions for the Christoffel symbols in spherical polar coordinates in $E_3$, and hence write down the differential equations satisfied by straight lines in $E_3$ in these coordinates.

2.19 In a four-dimensional space with coordinates $x^i$ ($i = 1, \ldots, 4$), if $L_j^i$ are independent of position and satisfy

$$\eta_{ij} = \sum_{h=1}^{4} \sum_{k=1}^{4} \eta_{hk} L_i^h L_j^k$$

where

$$(\eta_{ij}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

then the transformation

$$\bar{x}^i = \sum_{j=1}^{4} L_j^i x^j$$

is called a Lorentz transformation.

A set of $4^r$ quantities $T_{h_1 \cdots h_r}$ is said to constitute the components of a Lorentz tensor (or a 4-tensor) of rank $r$ if under a Lorentz transformation they transform according to the law

$$\bar{T}_{j_1 \cdots j_r} = \sum_{h_1=1}^{4} \cdots \sum_{h_r=1}^{4} L_{h_1}^{j_1} \cdots L_{h_r}^{j_r} T_{h_1 \cdots h_r}.$$

Show that if a Lorentz tensor vanishes in a given coordinate system, it vanishes in all other systems related to the first by a Lorentz transformation. Show that if $T_{ijk}$, $S_{ijk}$, and $V_{ij}$ are the components of Lorentz tensors of rank 3, 3, and 2, respectively, then $T_{ijk} + S_{ijk}$, $S_{ijk} V_{hl}$ and $\sum_{a=1}^{4} \sum_{b=1}^{4} \eta_{ab} V_{ab}$ are Lorentz tensors of rank 3, 5, and 0, respectively, whereas $\sum_{a=1}^{4} V_{aa}$ is not tensorial (Synge [3], Davis [1]).

2.20 In the notation of Problem 2.19 show that the set of all Lorentz transformations forms a group.

# 3

# TENSOR ANALYSIS
# ON MANIFOLDS

Since it is our objective to present a concrete and gradual development of the theory of tensors, introducing new concepts only when they become indispensable, the discussion of the previous chapters is carried out against the background of a Euclidean space $E_n$. The latter is characterized by the fact that it admits Cartesian coordinate systems, each of which covers $E_n$ completely. However, some of the most important applications of the tensor calculus are concerned with situations which require that the underlying space be of a far more general nature, such as a curved surface, which may not allow for the existence of a coordinate system by means of which it can be covered entirely. Thus we must now divest ourselves of the assumption that the underlying space is Euclidean; instead, it is now supposed that this role is played by a so-called differentiable manifold. The concept of a differentiable manifold is a somewhat abstract one: it can be described very roughly as an $n$-dimensional space $X_n$ which can be covered by open neighborhoods on each of which coordinate systems may be defined in such a manner as to ensure that pairs of such systems are related to each other by differentiable coordinate transformations. Smooth curves and surfaces in $E_3$ represent simple examples of differentiable manifolds.

The first section of this chapter is devoted to a very rudimentary description of such manifolds, which, it is hoped, is sufficient for all subsequent requirements. (A far more sophisticated and complete formulation is presented in the Appendix.) Again, as the result of further abstraction from the theory of tensors in $E_n$, tensor fields on $X_n$ are introduced, and it is immediately evident that the corresponding algebraic processes are formally identical with those described previously. However, when one turns to the calculus of tensors, one is immediately confronted with a formidable diffi-

culty which results directly from the fact that the derivatives of tensor components are not in general tensorial. Thus a new type of derivative, the so-called covariant derivative, is introduced for the purpose of differentiation of tensors. This, however, is possible solely on the strength of an additional assumption, namely, that the differentiable manifold is endowed with an affine connection.

The process of covariant differentiation forms the basis of the tensor calculus. In some aspects it is very similar to partial differentiation in the usual sense; nevertheless, a fundamentally important difference arises as an immediate consequence of the fact that the order of repeated covariant differentiation with respect to distinct coordinates $x^h$ and $x^k$ is by no means immaterial. In fact, this phenomenon leads directly to the concept of the curvature tensor of $X_n$ (and, already at this stage, it should be emphasized that this tensor vanishes identically in an $E_n$). A striking geometrical interpretation of this state of affairs is afforded by the behavior of parallel vector fields on $X_n$.

## 3.1  COORDINATE TRANSFORMATIONS ON DIFFERENTIABLE MANIFOLDS

In this section we shall give a very rough and intuitive description of the concept of a differentiable manifold. In order to attain an adequate degree of generality for our general theory we cannot restrict ourselves to spaces which can be covered completely by a single coordinate system (such as the $n$-dimensional Euclidean space $E_n$): simple examples of various kinds of surfaces embedded in $E_3$ indicate that, in general, no single coordinate system can exist which covers a given surface completely. This is true, for example, of the two-dimensional sphere in $E_3$, and consequently one considers certain regions on the sphere which are chosen such that it is possible to construct coordinate systems on each of these regions. This can be done in various ways; for instance, relative to a rectangular coordinate system $(x^1, x^2, x^3)$ of $E_3$, one may choose the hemisphere for which $x^1 > 0$, on which one can use $x^2$, $x^3$ as coordinates, for obviously this hemisphere can be mapped onto an open disk on the $(x^2, x^3)$-plane. Accordingly this hemisphere is referred to as a "coordinate neighborhood," and similarly the five other hemispheres corresponding to the restrictions $x^1 < 0, x^2 > 0, x^2 < 0, x^3 > 0,$ $x^3 < 0$ can be regarded as coordinate neighborhoods. These six hemispheres cover the sphere completely, and it is obviously feasible to use the coordinate system thus defined. In general, of course, the existence of suitable coordinate neighborhoods depends on the topological properties of the surface taken as a whole. We do not concern ourselves here with problems of this kind, the reader being referred to texts on differential topology. Instead, we simply

require that the underlying manifold with which we are concerned be a so-called differentiable manifold, so that the existence of coordinate neighborhoods with appropriate coordinate systems is guaranteed by definition.

An $n$-dimensional manifold is a point set $M$ which is covered completely by a countable set of neighborhoods $U_1, U_2, \ldots$, such that each point $P \in M$ belongs to at least one of these neighborhoods; it is assumed that a coordinate system is defined on each $U$ in the sense that one may assign in a unique manner $n$ real numbers $x^1, \ldots, x^n$ to each point $P \in U$. Thus as $P$ ranges over $U$, the corresponding numbers $x^1, \ldots, x^n$ range over an open domain $D$ of $E_n$; in other words, it is supposed that there exists a one-to-one mapping of each neighborhood $U$ onto $D$. This mapping is assumed to be continuous. The numbers $x^1, \ldots, x^n$ are called the coordinates of $P$; two distinct points of $U$ have distinct coordinates.

Let $U_1, U_2$ be *any* two coordinate neighborhoods on $E_n$ such that $U_1 \cap U_2$ is nonempty. Let $P \in U_1 \cap U_2$. Then, corresponding to the two coordinate systems on $U_1$ and $U_2$, we may assign to $P$ the two sets of coordinates $x^1, \ldots, x^n$ and $\bar{x}^1, \ldots, \bar{x}^n$, respectively, which, as before, we simply denote by $x^j$ and $\bar{x}^j$, it being understood that all Latin indices $j, h, k, \ldots$ range from 1 to $n$. (See Figure 5.) Clearly the values of $x^j$ and $\bar{x}^j$ must be somehow related: it will be assumed that this relation can be expressed in the form

$$\bar{x}^j = \bar{x}^j(x^h), \tag{1.1}$$

with inverse

$$x^h = x^h(\bar{x}^j), \tag{1.2}$$

(where we have used the compressed notation introduced in Section 2.4). The relations (1.1) and (1.2) represent coordinate transformations on the set $U_1 \cap U_2$. It will be supposed furthermore that the functions on the
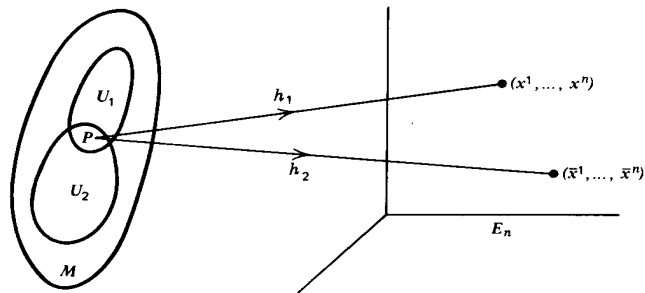


**Fig. 5**

right-hand sides of (1.1) and (1.2) are of class $C^\infty$ for all points $P \in U_1 \cap U_2$. If the manifold $M$ is such as to admit a construction of this kind, it is called an *n-dimensional differentiable manifold*; we shall henceforth denote such manifolds by $X_n$. (It should be emphasized that this is a very rough description of this concept: the precise formulation is considerably more sophisticated and is discussed more carefully in the Appendix.)

Because our definition implies the existence of the partial derivatives $\partial \bar{x}^j / \partial x^h$ and $\partial x^h / \partial \bar{x}^j$, we immediately infer the validity of the identities (2.4.29) and (2.4.30) in respect of the coordinate transformations (1.1), (1.2) on $X_n$. However, in order to simplify our subsequent notation, we shall now introduce the following convention.

SUMMATION CONVENTION.   When a lowercase Latin index such as $j, h, k, \ldots$ appears twice in a term, summation over that index is implied, the range of summation being $1, \ldots, n$. The letter $n$ is excluded from the summation convention; this letter will invariably denote the fixed dimension of the manifold under consideration.

Unless specifically stated otherwise, this convention is operative throughout (the reader may familiarize himself with this notation by applying it to the formulae of Section 2.4). In particular, we now write the identities (2.4.29) and (2.4.30), respectively, in the form

$$\delta_k^h = \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \bar{x}^j}{\partial x^k}, \tag{1.3}$$

and

$$\delta_i^j = \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^l}. \tag{1.4}$$

(Thus on the right-hand side of (1.3) summation over $j$ from 1 to $n$ is understood; similarly on the right-hand side of (1.4) a summation over $h$ is implied.)

From the product rule for determinants we immediately infer from (1.3) that

$$\frac{\partial(x^1, \ldots, x^n)}{\partial(\bar{x}^1, \ldots, \bar{x}^n)} \cdot \frac{\partial(\bar{x}^1, \ldots, \bar{x}^n)}{\partial(x^1, \ldots, x^n)} = 1, \tag{1.5}$$

in terms of the notation (2.4.2). Thus the Jacobian of the transformation (1.2), namely,

$$J = \frac{\partial(x^1, \ldots, x^n)}{\partial(\bar{x}^1, \ldots, \bar{x}^n)}, \tag{1.6}$$

does not vanish on $U_1 \cap U_2$: this is a direct consequence of our definition.

The manifold $X_n$ is said to be *orientable* if it is possible to choose coordinate neighborhoods and corresponding coordinate systems such that each of these Jacobians is positive (on the respective intersections of the coordinate neighborhoods). In conclusion we recall that a real-valued function $f$ of the $n$ variables $x^j$ is said to be of class $C^r$ at a point $P$ of a coordinate neighborhood $U$ if it possesses continuous derivatives with respect to each of $x^1, \ldots, x^n$ up to and including the $r$th order at $P$. It should be observed that, by virtue of our construction, the property of being of class $C^r$ on $X_n$ is independent of the particular choice of coordinates.

## 3.2  TENSOR ALGEBRA ON MANIFOLDS

We now present the most general definition of a tensor; since we have already dealt with several special cases under different conditions, we proceed in a purely formal manner without further motivation.

### DEFINITION

*A set of $n^{r+s}$ quantities $T^{h_1 \cdots h_r}_{k_1 \cdots k_s}$ is said to constitute the components of a tensor of type $(r, s)$ at a point $P$ of a differentiable manifold $X_n$, if, under the coordinate transformation* (1.1), *these quantities transform according to the law*

$$\overline{T}^{j_1 \cdots j_r}_{l_1 \cdots l_s} = \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \cdots \frac{\partial \bar{x}^{j_r}}{\partial x^{h_r}} \frac{\partial x^{k_1}}{\partial \bar{x}^{l_1}} \cdots \frac{\partial x^{k_s}}{\partial \bar{x}^{l_s}} T^{h_1 \cdots h_r}_{k_1 \cdots k_s}. \tag{2.1}$$

The *rank* of $T$ is $r + s$.

Again, our definition contains an explicit reference to the point $P$ at which the tensor is defined because the coordinates $x^k$ of that point enter as arguments in the coefficients $\partial \bar{x}^j / \partial x^h$ on the right-hand side of (2.1). Strictly speaking, we should have written these derivatives in the form $\partial \bar{x}^j(x^k)/\partial x^h$ in order to indicate explicitly that they are to be evaluated at $P$.

Because of (1.4) the transformation law (2.1) possesses an inverse, namely,

$$T^{h_1 \cdots h_r}_{k_1 \cdots k_s} = \frac{\partial x^{h_1}}{\partial \bar{x}^{j_1}} \cdots \frac{\partial x^{h_r}}{\partial \bar{x}^{j_r}} \frac{\partial \bar{x}^{l_1}}{\partial x^{k_1}} \cdots \frac{\partial \bar{x}^{l_s}}{\partial x^{k_s}} \overline{T}^{j_1 \cdots j_r}_{l_1 \cdots l_s}, \tag{2.2}$$

which, in turn, implies (2.1).

Once more it should be emphasized that, *if all the components of any tensor vanish in a given coordinate system, they will vanish in any other system.* As special cases of the above definition we have *contravariant vectors* as type $(1, 0)$ tensors, whose transformation law is given by

$$\overline{A}^j = \frac{\partial \bar{x}^j}{\partial x^h} A^h; \tag{2.3}$$

similarly, *covariant vectors* are type $(0, 1)$ tensors, whose transformation law is

$$\overline{C}_j = \frac{\partial x^h}{\partial \bar{x}^j} C_h, \tag{2.4}$$

while type $(1, 1)$ tensors satisfy the transformation law

$$\overline{T}^j_i = \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^i} T^h_k. \tag{2.5}$$

In particular, we can write (1.4) in the form

$$\delta^j_i = \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \bar{x}^i} \delta^h_k, \tag{2.6}$$

showing that the Kronecker delta is a type $(1, 1)$ tensor.

A type $(0, 0)$ tensor is simply a scalar or invariant.

We now give a formal treatment of algebraic processes which may be applied to tensors at a fixed point $P$ of $X_n$.

### Addition

Let $S^{h_1 \cdots h_r}_{k_1 \cdots k_s}$ be a type $(r, s)$ tensor defined at $P$. Its transformation law, according to (2.1), is given by

$$\overline{S}^{j_1 \cdots j_r}_{l_1 \cdots l_s} = \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \cdots \frac{\partial \bar{x}^{j_r}}{\partial x^{h_r}} \cdot \frac{\partial x^{k_1}}{\partial \bar{x}^{l_1}} \cdots \frac{\partial x^{k_s}}{\partial \bar{x}^{l_s}} S^{h_1 \cdots h_r}_{k_1 \cdots k_s}.$$

If this is added to (2.1), we obtain

$$\overline{T}^{j_1 \cdots j_r}_{l_1 \cdots l_s} + \overline{S}^{j_1 \cdots j_r}_{l_1 \cdots l_s}$$
$$= \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \cdots \frac{\partial \bar{x}^{j_r}}{\partial x^{h_r}} \frac{\partial x^{k_1}}{\partial \bar{x}^{l_1}} \cdots \frac{\partial x^{k_s}}{\partial \bar{x}^{l_s}} (T^{h_1 \cdots h_r}_{k_1 \cdots k_s} + S^{h_1 \cdots h_r}_{k_1 \cdots k_s}), \tag{2.7}$$

which shows that the sums $T^{h_1 \cdots h_r}_{k_1 \cdots k_s} + S^{h_1 \cdots h_r}_{k_1 \cdots k_s}$ are components of a tensor of type $(r, s)$ at $P$. Also, it follows directly from (2.1) that the multiplication by a scalar or a constant of the components of a tensor of type $(r, s)$ yields a tensor of the same type. *Thus the set of all tensors of type $(r, s)$ at the point $P$ of $X_n$ constitutes a vector space of dimension $n^{r+s}$.* In particular, the set of all contravariant vectors at $P$ defines the so-called $n$-dimensional *tangent space* $T_n(P)$, while the set of all covariant vectors at $P$ constitutes the *dual tangent space* $T_n^*(P)$, which is also of dimension $n$.

We remark that it is not possible to add tensors whose types are distinct, nor is it permissible to add tensors defined at different points $P$ and $Q$ of $X_n$.

## Multiplication

It is always possible to multiply tensors of arbitrary type at the point $P$ of $X_n$ componentwise. More precisely, *the multiplication of the components of two tensors of types $(r_1, s_1)$ and $(r_2, s_2)$ at $P$ yields a tensor of type $(r_1 + r_2, s_1 + s_2)$ at $P$*. Rather than prove this statement in its fullest generality, which would involve a profusion of indices and subscripts thereof, let us consider a type $(2, 1)$ tensor and a type $(0, 2)$ tensor, whose transformation laws are respectively given by

$$\overline{T}^{jl}{}_m = \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial \overline{x}^l}{\partial x^k} \frac{\partial x^p}{\partial \overline{x}^m} T^{hk}{}_p, \tag{2.8}$$

and

$$\overline{S}_{qr} = \frac{\partial x^u}{\partial \overline{x}^q} \frac{\partial x^v}{\partial \overline{x}^r} S_{uv}. \tag{2.9}$$

The products of the components of these tensors transform according to

$$\overline{T}^{jl}{}_m \overline{S}_{qr} = \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial \overline{x}^l}{\partial x^k} \frac{\partial x^p}{\partial \overline{x}^m} \frac{\partial x^u}{\partial \overline{x}^q} \frac{\partial x^v}{\partial \overline{x}^r} T^{hk}{}_p S_{uv}, \tag{2.10}$$

which is obviously the transformation law of the components

$$V^{hk}{}_{puv} = T^{hk}{}_p S_{uv} \tag{2.11}$$

of a tensor of type $(2, 3)$.

This process of multiplication can be combined with the process of addition of tensors, provided that their respective types are appropriate. Clearly the commutative, associative, and distributive laws are satisfied.

## Contraction

Given a tensor of type $(r, s)$, one may select a pair of indices, of which one is a superscript, the other being a subscript, and replace them by two identical indices, summation over the latter being implied by virtue of the summation convention. This process is known as contraction, and *the quantities obtained by contraction constitute the components of a tensor of type $(r - 1, s - 1)$*. Again, we shall not prove this assertion generally; instead, we shall verify it for the case of a type $(2, 1)$ tensor. Thus in $(2.8)$, let us identify the indices $l$ and $m$, denoting both by $q$, which gives

$$\overline{T}^{jq}{}_q = \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial \overline{x}^q}{\partial x^k} \frac{\partial x^p}{\partial \overline{x}^q} T^{hk}{}_p,$$

it being understood that a summation over $q$ is implied on both sides. Then,

because of $(1.3)$, the product $(\partial \overline{x}^q / \partial x^k)(\partial x^p / \partial \overline{x}^q)$ on the right-hand side is merely the Kronecker delta $\delta_k^p$, and accordingly we obtain

$$\overline{T}^{jq}{}_q = \frac{\partial \overline{x}^j}{\partial x^h} T^{hp}{}_p, \tag{2.12}$$

which is the transformation law of a type $(1, 0)$ tensor.

Clearly the process of contraction of a type $(1, 1)$ tensor yields a scalar; in particular, for the case of the Kronecker delta we have

$$\delta_j^j = \delta_1^1 + \cdots + \delta_n^n = n. \tag{2.13}$$

Furthermore, one may form the products of components of tensors of arbitrary types and then contract (provided, of course, that the process of multiplication yields a type $(r, s)$ tensor with $r \geq 1$ and $s \geq 1$). For instance, if the components $T^{hk}$ of a type $(2, 0)$ tensor which satisfies

$$\overline{T}^{jl} = \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial \overline{x}^l}{\partial x^k} T^{hk},$$

are multiplied by the components $C_p$, $F_q$ of two covariant vectors, we obtain a type $(2, 2)$ tensor, whose transformation law is given by

$$\overline{T}^{jl} \overline{C}_p \overline{F}_q = \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial \overline{x}^l}{\partial x^k} \frac{\partial x^r}{\partial \overline{x}^p} \frac{\partial x^s}{\partial \overline{x}^q} T^{hk} C_r F_s;$$

if we contract over the indices $l$, $p$, this reduces to the transformation law of a type $(1, 1)$ tensor, and if we contract once more over the remaining indices $j$, $q$ we obtain the scalar

$$\overline{T}^{jl} \overline{C}_l \overline{F}_j = T^{hk} C_k F_h.$$

## Symmetrization

A tensor is said to be symmetric in a pair of superscripts (or in a pair of subscripts) if an interchange of the indices concerned does not affect the values of the components of that tensor. If, on the other hand, this process is tantamount to the multiplication of each component by $-1$, the tensor is said to be skew-symmetric (or anti-symmetric) in these indices. For instance, if $A_{hj}$, $B_{hj}$ denote the components of two type $(0, 2)$ tensors, and if

$$A_{hj} = A_{jh}, \tag{2.14}$$

while

$$B_{hj} = -B_{jh}, \tag{2.15}$$

the first is a symmetric tensor, while the second is skew-symmetric. Clearly the equations $(2.14)$ and $(2.15)$ will hold in any coordinate system as a direct

consequence of the form of the transformation laws of tensors: accordingly *all symmetry and skew-symmetry properties of tensors are independent of the choice of the coordinate system.*

Given any tensor of type $(r, s)$ with $r > 1$, or $s > 1$, one can always construct from it a symmetric and a skew-symmetric tensor in any pair of subscripts, or any pair of superscripts. For example, in the case of a type $(0, 2)$ tensor $C_{hj}$ we can define

$$S_{hj} = \tfrac{1}{2}(C_{hj} + C_{jh}), \tag{2.16}$$

and

$$T_{hj} = \tfrac{1}{2}(C_{hj} - C_{jh}), \tag{2.17}$$

these being respectively symmetric and skew-symmetric. Naturally these processes may be applied to tensors of arbitrary types ($r > 1$ or $s > 1$), the remaining indices being held fixed. The process (2.16) is often referred to as *symmetrization.* Furthermore, any such tensor can be written as the sum of its so-called symmetric and skew-symmetric parts; for instance, one obviously has

$$C_{hj} = \tfrac{1}{2}(C_{hj} + C_{jh}) + \tfrac{1}{2}(C_{hj} - C_{jh}). \tag{2.18}$$

*Remark.* In the literature $S_{hj}$ and $T_{hj}$ as defined by (2.16) and (2.17) are often respectively denoted by $C_{(hj)}$ and $C_{[hj]}$. This notation will not be used here. Also, the process of symmetrization and its converse can be generalized to an operation involving more than two indices: however, we shall defer the discussion of this technique until we have dealt with the so-called numerical tensors.

### Criteria for Tensor Character: the Quotient Theorems

It was seen above that the product of two tensors yields another tensor. This immediately raises the question as to a possible converse: if a tensor, multiplied by certain components, is known to give rise to another tensor, can one conclude that these components themselves constitute a tensor? Sometimes the answer to this is in the affirmative, and may therefore represent a criterion for the tensor character of given components: criteria of this kind are contained in the so-called *quotient theorems*. It should be emphasized, however, that these theorems should be applied with utmost caution, as will be evident from our remarks below.

As a first example of such a theorem, let us consider the following state of affairs. Suppose that, at a fixed point $P$ of $X_n$, we are given a set of $n$ quantities $a_h$ subject to the understanding that $a_h X^h$ is a scalar for *any* contravariant vector $X^h$ at $P$, so that we may write

$$a_h X^h = \phi, \tag{2.19}$$

in which $\phi$ denotes a scalar (which depends, of course, on $a_h$ and $X^h$). Does this imply that the $a_h$ constitute the components of a vector at $P$? We shall see that this is indeed the case. Denoting the as yet unknown transform of $a_h$ under (1.1) by $\bar{a}_j$, we know that in the $\bar{x}$-coordinate system the condition (2.19) reads

$$\bar{a}_j \bar{X}^j = \bar{\phi}. \tag{2.20}$$

Since $\phi$ is a scalar, we have $\bar{\phi} = \phi$. Furthermore, since $X^h$ is a contravariant vector, it follows that

$$\bar{X}^j = \frac{\partial \bar{x}^j}{\partial x^h} X^h; \tag{2.21}$$

accordingly subtraction of (2.20) from (2.19) gives

$$\left(a_h - \bar{a}_j \frac{\partial \bar{x}^j}{\partial x^h}\right) X^h = 0. \tag{2.22}$$

Here we note that a summation over $h$ is implied, so that we cannot assert directly that the coefficients of $X^h$ vanish. However, we now invoke the additional hypothesis according to which (2.19) and hence (2.22) are valid for *any* $X^h$. We may therefore choose a vector $X^h$ at $P$ with components $(1, 0, \ldots, 0)$. Equation (2.22) then reduces to

$$a_1 = \frac{\partial \bar{x}^j}{\partial x^1} \bar{a}_j.$$

Similarly, choosing a vector with components $(0, 1, \ldots, 0)$ at $P$, we infer that

$$a_2 = \frac{\partial \bar{x}^j}{\partial x^2} \bar{a}_j.$$

Continuing in this manner, we find that

$$a_h = \frac{\partial \bar{x}^j}{\partial x^h} \bar{a}_j,$$

which is the transformation law for covariant vectors. We have therefore established that if $a_h X^h$ *is a scalar for any vector* $X^h$ *at* $P$, *then the* $a_h$ *constitute the components of a covariant vector at* $P$. This result is called a quotient theorem; however, in its application one must be absolutely certain that the contravariant vector $X^h$ which appears in its enunciation is *arbitrary*, and this hypothesis represents a very strong condition which is not often satisfied.

As a second important case suppose that we are given a set of $n^2$ quantities $a_{hk}$ of which it is known that $a_{hk} X^h X^k$ is a scalar $\psi$ for *any* vector $X^h$ at $P$.

Again, can it be inferred that the $a_{hk}$ constitute the components of a type $(0, 2)$ tensor? We shall now see that this is *not* true in general. By hypothesis we have

$$a_{hk} X^h X^k = \psi$$

in the given coordinate system, and similarly

$$\bar{a}_{jl} \bar{X}^j \bar{X}^l = \bar{\psi}$$

in the $\bar{x}$-coordinate system, in which we have denoted the as yet unknown transforms of $a_{hk}$ by $\bar{a}_{jl}$. Using (2.21), and the fact that $\bar{\psi} = \psi$, we obtain by subtraction

$$\left( a_{hk} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^l}{\partial x^k} \right) X^h X^k = 0. \tag{2.23}$$

Because a summation is implied over $h$ and $k$, we cannot infer directly that the coefficients of $X^h X^k$ vanish. Again, by successively choosing vectors $X^j$ at $P$ with components $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$, we deduce from (2.23) that

$$a_{11} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^l}{\partial x^1} = 0, \qquad a_{22} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^2} \frac{\partial \bar{x}^l}{\partial x^2} = 0, \qquad \ldots, \qquad \text{etc.,} \tag{2.24}$$

but this tells us nothing about the terms involving $a_{jh}$ with $j \neq h$. Accordingly we now choose a vector $X^j$ at $P$ with components $(X^1, X^2, 0, \ldots, 0)$. For this particular vector (2.23) becomes

$$\left( a_{11} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^l}{\partial x^1} \right) X^1 X^1 + \left( a_{12} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^l}{\partial x^2} \right) X^1 X^2$$

$$+ \left( a_{21} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^2} \frac{\partial \bar{x}^l}{\partial x^1} \right) X^2 X^1 + \left( a_{22} - \bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^2} \frac{\partial \bar{x}^l}{\partial x^2} \right) X^2 X^2 = 0.$$

Because of (2.24) the coefficients of $X^1 X^1$ and $X^2 X^2$ vanish; also, since

$$\bar{a}_{jl} \frac{\partial \bar{x}^j}{\partial x^2} \frac{\partial \bar{x}^l}{\partial x^1} = \bar{a}_{lj} \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^l}{\partial x^2}$$

by a mere relabeling of the indices $j$ and $l$, it is found that

$$\left[ (a_{12} + a_{21}) - (\bar{a}_{jl} + \bar{a}_{lj}) \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^l}{\partial x^2} \right] X^1 X^2 = 0.$$

Thus, choosing $X^1 = 1$, $X^2 = 1$, it is seen that

$$a_{12} + a_{21} = (\bar{a}_{jl} + \bar{a}_{lj}) \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^l}{\partial x^2}.$$

Again, this process may be repeated to yield

$$a_{hk} + a_{kh} = (\bar{a}_{jl} + \bar{a}_{lj}) \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^l}{\partial x^k}. \tag{2.25}$$

This is, in fact, the transformation law of a type $(0, 2)$ tensor, but it refers to $a_{hk} + a_{kh}$, that is, the symmetric part of $2a_{hk}$, and *not* to $a_{hk}$ as such. Indeed, nothing can be inferred about the tensorial character of the skew-symmetric part of $a_{hk}$ from our hypothesis. This is quite natural, for the skew-symmetric part of $a_{hk}$ contributes nothing to the scalar $\psi$. This is so because

$$(a_{hk} - a_{kh}) X^h X^k = a_{hk} X^h X^k - a_{kh} X^h X^k = a_{hk} X^h X^k - a_{hk} X^k X^h = 0$$

identically, where, in the last step, we have interchanged the indices $h$ and $k$.

Accordingly, the quotient theorem for this case must be stated as follows: *If a set of $n^2$ quantities $a_{hk}$ is such that for any contravariant vector $X^h$ at a point $P$ of $X_n$ the sum $a_{hk} X^h X^k$ is a scalar, then the symmetric parts $\frac{1}{2}(a_{hk} + a_{kh})$ of $a_{hk}$ are the components of a type $(0, 2)$ tensor.*

As a corollary we note that, if in addition to the above hypothesis, we are given that the $a_{hk}$ are symmetric, then it may be concluded that the $a_{hk}$ are the components of a type $(0, 2)$ tensor.

Clearly other quotient theorems can be established for tensors of different types. In each case the hypotheses of these theorems are extremely stringent and should be examined with utmost care before the theorems are applied. Their incorrect application has led to numerous faulty arguments in the literature; as a rule it is best to avoid them and to derive the transformation properties of given quantities directly by ad hoc methods.

## 3.3   TENSOR FIELDS AND THEIR DERIVATIVES

In our discussion of tensor algebra it was repeatedly stressed that all algebraic operations are concerned with tensors defined at one and the same point $P$ of our differentiable manifold $X_n$. Indeed, the definition of a tensor as given above is formulated solely in terms of components prescribed at $P$. In many applications, however, tensorial quantities are defined not only at a point but on a finite region of $X_n$, and the restriction to a single point is unrealistic. It may be necessary to compare, say, contravariant vectors at distinct points of $X_n$, and our algebraic apparatus is totally inadequate for purposes of this kind.

Examples of tensorial quantities defined on finite regions are plentiful. For instance, if one has a differentiable scalar function $\phi(x^h)$ defined on a region, the components $\partial \phi / \partial x^h$ of its gradient are also defined at each point of that region, and accordingly one speaks of a gradient *field*. Similarly the vectors which describe the strength of an electromagnetic field on a region of $E_3$ represent a vector field.

In order to be able to extend the general tensor concept in this sense, we formulate the definition of a tensor field as follows, noting that it is now necessary to regard the components of a tensor as functions of the coordinates $x^q$ of the points of the region of $X_n$ on which these components are defined.

## DEFINITION

*A set of $n^{r+s}$ functions $T^{h_1\cdots h_r}{}_{k_1\cdots k_s}(x^q)$ is said to constitute the components of a tensor field of type $(r, s)$ on the manifold $X_n$, if, under the coordinate transformation (1.1), these functions transform according to the relation*

$$\bar{T}^{j_1\cdots j_r}{}_{l_1\cdots l_s}(\bar{x}^p) = \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \cdots \frac{\partial \bar{x}^{j_r}}{\partial x^{h_r}} \cdot \frac{\partial x^{k_1}}{\partial \bar{x}^{l_1}} \cdots \frac{\partial x^{k_s}}{\partial \bar{x}^{l_s}} T^{h_1\cdots h_r}{}_{k_1\cdots k_s}(x^q). \tag{3.1}$$

In this formulation we have explicitly indicated the dependence of the components on the positional coordinates, namely, by writing $T^{\cdots}{}_{\cdots}(x^q)$ on the right-hand side, and $\bar{T}^{\cdots}{}_{\cdots}(\bar{x}^p)$ on the left; here $x^q$ refers to the coordinates $x^1, \ldots, x^n$ of any point of $X_n$ at which $T^{\cdots}{}_{\cdots}$ is defined, while $\bar{x}^p$ refers to the coordinates $\bar{x}^1, \ldots, \bar{x}^n$ *of the same point* in the $\bar{x}$-system. Clearly these are also the arguments which respectively enter the coefficients $\partial \bar{x}^j/\partial x^h$ and $\partial x^k/\partial \bar{x}^l$ in (3.1).

In order to be able to relate tensors at distinct points it is manifestly necessary to use a process of differentiation. A tensor is said to be of class $C^p$ on a region of $X_n$ if its components are class $C^p$ functions of the coordinates: obviously this requirement is independent of the choice of the coordinate system. Thus for a tensor field of class $C^p(p \geq 1)$ we can form the partial derivatives of its components, and we now investigate how these partial derivatives may be used. We begin by examining a few simple cases.

First, let us consider a differentiable scalar field $\phi(x^k)$. As we pass from a point $P$ with coordinates $x^k$ to a neighboring point $Q$ with coordinates $(x^k + dx^k)$, the corresponding differential is given by

$$d\phi = \frac{\partial \phi}{\partial x^k} dx^k. \tag{3.2}$$

Recalling that the components $\partial \phi/\partial x^k$ of the gradient of $\phi$ constitute a covariant vector, while the $dx^k$ represent the components of a contravariant vector by virtue of the fact that for any such displacement $d\bar{x}^j = (\partial \bar{x}^j/\partial x^k) dx^k$, we see that the right-hand side of (3.2) is a scalar (in fact, it is an inner product). Thus $d\phi$ is a scalar: clearly no difficulties are encountered here.

However, when we consider a second example, namely the case of a class $C^1$ contravariant vector field $X^h(x^k)$, we shall see that certain difficulties

emerge almost immediately. Again, corresponding to the transition from $P$ to $Q$, the differential $dX^h$ of $X^h$ is given by

$$dX^h = \frac{\partial X^h}{\partial x^k} dx^k. \tag{3.3}$$

These differentials would represent the components of a contravariant vector if the partial derivatives $\partial X^h/\partial x^k$ were the components of a type $(1, 1)$ tensor, but generally this is not true.

This is seen directly by an examination of the transformation law of our vector field, namely

$$\bar{X}^j(\bar{x}^p) = \frac{\partial \bar{x}^j}{\partial x^h} X^h(x^q), \tag{3.4}$$

which is to be differentiated with respect to $\bar{x}^k$. In doing so, we note that the coefficients $\partial \bar{x}^j/\partial x^h$ are given as functions of $x^l$, so that the chain rule must be used; a similar remark applies also to $X^h$. Accordingly it is found that

$$\frac{\partial \bar{X}^j}{\partial \bar{x}^k} = \frac{\partial^2 \bar{x}^j}{\partial x^l \partial x^h} \frac{\partial x^l}{\partial \bar{x}^k} X^h + \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^l}{\partial \bar{x}^k} \frac{\partial X^h}{\partial x^l}. \tag{3.5}$$

From this relation it is quite evident that the partial derivatives $\partial X^h/\partial x^l$ are *not* the components of a tensor: the presence of the first term on the right-hand side of (3.5) is the reason for this state of affairs. [This term is due to the fact that the coordinate transformation (1.1) is not assumed to be linear. If one restricts oneself to linear transformations, one has $\partial^2 \bar{x}^j/\partial x^l \partial x^h = 0$; thus in the case of an *affine* vector field $X^j(x^h)$ the partial derivatives (3.5) represent the components of a type $(1, 1)$ affine tensor.]

If we multiply (3.5) by $d\bar{x}^k$ and sum over $k$, noting that

$$\frac{\partial x^l}{\partial \bar{x}^k} d\bar{x}^k = dx^l, \tag{3.6}$$

we find that

$$\frac{\partial \bar{X}^j}{\partial \bar{x}^k} d\bar{x}^k = \frac{\partial^2 \bar{x}^j}{\partial x^l \partial x^h} X^h dx^l + \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial X^h}{\partial x^l} dx^l,$$

or, using (3.3) and its counterpart in the $\bar{x}$-system:

$$d\bar{X}^j = \frac{\partial \bar{x}^j}{\partial x^h} dX^h + \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} X^h dx^k. \tag{3.7}$$

This, then, is the transformation law of the differentials of the contravariant vector field $X^h(x^k)$: we reiterate that these differentials are not the components of a tensorial function. (It should be remarked that (3.7) can

be obtained more directly by considering the differentials of (3.4); however, the present method is used since (3.5) will be required later.)

Because of this defect we must now endeavor to construct a new type of differential which has some of the properties generally associated with the usual concept of a differential, but which is also tensorial. In our attempt to do so we shall be guided by the analysis of Section 2.5 concerning the derivatives of vectors relative to a curvilinear coordinate system in $E_n$; we saw that it was necessary to augment these derivatives by an additional term involving the Christoffel symbols in order to obtain a tensorial expression. Under the present circumstances we are not in possession of quantities of this kind a priori, and we shall therefore have to endow our manifold with suitable functions which play a role analogous to that of the Christoffel symbols. This is to be done as follows.

Let us denote the new type of differential which we are seeking by $DX^j$, it being assumed in view of the analogy referred to above that $DX^j$ is of the form

$$DX^j = dX^j + P^j(x^p, X^h, dx^k). \tag{3.8}$$

Here the components $P^j$ are to be specified by means of certain fairly natural conditions which are to be imposed on $DX^j$.

We observe that, for any two contravariant vector fields $X^j$, $Z^j$, the ordinary differential satisfies the condition $d(X^j + Z^j) = dX^j + dZ^j$, while it is evident from (3.3) that $dX^j$ is linear in $dx^k$. Thus the following analogous algebraic requirements are imposed upon the operator $D$:

1.   For any two contravariant vector fields $X^j$ and $Z^j$, $D(X^j + Z^j) = DX^j + DZ^j$, which is satisfied if $P^j$ is linear in $X^j$.
2.   $DX^j$ is linear in $dx^k$.

Accordingly the components $P^j$ on the right-hand side of (3.8) must be linear homogeneous functions of $X^h$ and $dx^k$; they are thus expressible in the form

$$P^j(x^p, X^h, dx^k) = \Gamma_h{}^j{}_k(x^p)X^h \, dx^k, \tag{3.9}$$

in which the coefficients $\Gamma_h{}^j{}_k(x^p)$ depend solely on the positional coordinates $x^p$. (Although we have endowed these coefficients with two subscripts and a superscript, it is not implied that the $\Gamma_h{}^j{}_k$ are the components of a tensor.)

Finally, it is stipulated that

3.   The $DX^j$ constitute the components of a contravariant vector:

$$D\bar{X}^j = \frac{\partial \bar{x}^j}{\partial x^h} DX^h. \tag{3.10}$$

From (3.8), and its counterpart in the "barred" coordinate system, we therefore infer that

$$d\bar{X}^j + \bar{P}^j = \frac{\partial \bar{x}^j}{\partial x^h} dX^h + \frac{\partial \bar{x}^j}{\partial x^l} P^l. \tag{3.11}$$

In order to obtain the resulting transformation law for the components $P^j$, we subtract (3.7) from (3.11), which gives

$$\bar{P}^j = \frac{\partial \bar{x}^j}{\partial x^l} P^l - \frac{\partial^2 \bar{x}^j}{\partial x^h \, \partial x^k} X^h \, dx^k. \tag{3.12}$$

Clearly this is a necessary and sufficient condition for the validity of (3.10); however, for our future requirements it is desirable that we derive the corresponding transformation law for the coefficients $\Gamma_h{}^j{}_k$ which appear in (3.9). By means of (3.4) and the inverse of (3.6) we can write the counterpart of (3.9) in the barred coordinate system as follows:

$$\bar{P}^j = \bar{\Gamma}_m{}^j{}_p \bar{X}^m \, d\bar{x}^p = \bar{\Gamma}_m{}^j{}_p \frac{\partial \bar{x}^m}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} X^h \, dx^k. \tag{3.13}$$

This, together with (3.9), is substituted in (3.12) to yield

$$\bar{\Gamma}_m{}^j{}_p \frac{\partial \bar{x}^m}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} X^h \, dx^k = \left( \frac{\partial \bar{x}^j}{\partial x^l} \Gamma_h{}^l{}_k - \frac{\partial^2 \bar{x}^j}{\partial x^h \, \partial x^k} \right) X^h \, dx^k. \tag{3.14}$$

In order that this relation be satisfied it is sufficient that the transformation law of the $\Gamma_h{}^l{}_k$ is given by

$$\bar{\Gamma}_m{}^j{}_p \frac{\partial \bar{x}^m}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} = \frac{\partial \bar{x}^j}{\partial x^l} \Gamma_h{}^l{}_k - \frac{\partial^2 \bar{x}^j}{\partial x^h \, \partial x^k}. \tag{3.15}$$

Furthermore, this condition is also necessary by virtue of the fact that the components $X^h$, $dx^k$ which appear in (3.14) are entirely arbitrary. Also, we can solve (3.15) for $\bar{\Gamma}_m{}^j{}_p$ in terms of $\Gamma_h{}^l{}_k$ by multiplying by $\partial x^h/\partial \bar{x}^r$, $\partial x^k/\partial \bar{x}^s$, which gives

$$\bar{\Gamma}_m{}^j{}_p \left( \frac{\partial \bar{x}^m}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^r} \right) \left( \frac{\partial \bar{x}^p}{\partial x^k} \frac{\partial x^k}{\partial \bar{x}^s} \right) = \frac{\partial \bar{x}^j}{\partial x^l} \Gamma_h{}^l{}_k \frac{\partial x^h}{\partial \bar{x}^r} \frac{\partial x^k}{\partial \bar{x}^s} - \frac{\partial^2 \bar{x}^j}{\partial x^h \, \partial x^k} \frac{\partial x^h}{\partial \bar{x}^r} \frac{\partial x^k}{\partial \bar{x}^s}.$$

We now apply (1.4) to the left-hand side above, obtaining $\bar{\Gamma}_m{}^j{}_p \delta_r^m \delta_s^p = \bar{\Gamma}_r{}^j{}_s$, and in the resulting relation we replace the indices $r$, $s$ by $m$, $p$ respectively. It is thus found that

$$\bar{\Gamma}_m{}^j{}_p = \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p} \Gamma_h{}^l{}_k - \frac{\partial^2 \bar{x}^j}{\partial x^h \, \partial x^k} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p}. \tag{3.16}$$

Any set of three-index symbols $\Gamma_h{}^j{}_k$ whose transformation law is given by (3.16) is said to constitute the *components of an affine connection, or connection coefficients*, on our differentiable manifold $X_n$. In terms of this connection we may now, by means of (3.8) and (3.9), write

$$DX^j = dX^j + \Gamma_h{}^j{}_k X^h \, dx^k, \tag{3.17}$$

which is known as the *absolute* or *covariant differential* of the vector field $X^j$: this was the object of our search.

*Remark 1.* In a certain sense the latter terminology is unfortunate as it involves the term "covariant" in a manner which is inconsistent with its previous use. Here the adjective covariant is used to indicate that (3.17) is tensorial.

*Remark 2.* Thus far we have merely constructed the absolute differential of a contravariant vector field: we shall show, however, that similar processes can be applied to tensors of arbitrary type.

*Remark 3.* It should be reiterated that, despite the notation, the coefficients $\Gamma_h{}^j{}_k$ of the affine connection are *not* the components of a tensor.

*Remark 4.* Obviously the conditions 1–3 do not completely specify the $\Gamma_h{}^j{}_k$. In fact, any set of $n^3$ functions of position which transform according to (3.16) defines an affine connection by means of which absolute differentials may be constructed. Thus, before one can construct such differentials, the differentiable manifold $X_n$ must be endowed with these coefficients. The manifold is then called an *affinely connected space* (see Schrödinger [1]).

*Remark 5.* The transformation law (3.16) can be expressed in a slightly different form as follows. We recall that, according to (1.4),

$$\frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^m} = \delta_m^j. \tag{3.18}$$

This is differentiated partially with respect to $x^k$, it being recalled that $\partial x^h / \partial \bar{x}^m$ is given as a function of $\bar{x}^q$, so that the chain rule has to be applied to this term. It is seen that

$$\frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial x^h}{\partial \bar{x}^m} = -\frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 x^h}{\partial \bar{x}^m \partial \bar{x}^q} \frac{\partial \bar{x}^q}{\partial x^k}, \tag{3.19}$$

which is multiplied by $\partial x^k / \partial \bar{x}^p$ to give

$$\frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p} = -\frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 x^h}{\partial \bar{x}^m \partial \bar{x}^q} \frac{\partial \bar{x}^q}{\partial x^k} \frac{\partial x^k}{\partial \bar{x}^p}$$

$$= -\frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 x^h}{\partial \bar{x}^m \partial \bar{x}^q} \delta_p^q = -\frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 x^h}{\partial \bar{x}^m \partial \bar{x}^p}. \tag{3.20}$$

This is substituted in (3.16), which then becomes

$$\bar{\Gamma}_m{}^j{}_p = \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p} \Gamma_h{}^l{}_k + \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 x^h}{\partial \bar{x}^m \partial \bar{x}^p}. \tag{3.21}$$

An equivalent form of this is

$$\frac{\partial^2 x^h}{\partial \bar{x}^m \partial \bar{x}^p} = \frac{\partial x^h}{\partial \bar{x}^j} \bar{\Gamma}_m{}^j{}_p - \frac{\partial x^l}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p} \Gamma_l{}^h{}_k, \tag{3.22}$$

which results from (3.21) by a repeated application of (1.4) and a change of indices.

With the aid of the formalism developed above, it is now a simple matter to define an appropriate absolute differential of a covariant vector field $Y_h = Y_h(x^k)$. We recall that the transformation law of such a field is given by

$$Y_h = \frac{\partial \bar{x}^j}{\partial x^h} \bar{Y}_j, \tag{3.23}$$

so that the differentials of $Y_h$ relative to the two coordinate systems are related by

$$dY_h = \frac{\partial \bar{x}^j}{\partial x^h} d\bar{Y}_j + \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \bar{Y}_j \, dx^k. \tag{3.24}$$

We can eliminate the second term on the right-hand side by means of (3.15), which is multiplied by $\bar{Y}_j \, dx^k$ for this purpose, giving

$$\frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \bar{Y}_j \, dx^k = \Gamma_h{}^l{}_k \, dx^k \bar{Y}_j \frac{\partial \bar{x}^j}{\partial x^l} - \bar{Y}_j \bar{\Gamma}_m{}^j{}_p \frac{\partial \bar{x}^m}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} \, dx^k$$

$$= \Gamma_h{}^l{}_k Y_l \, dx^k - \frac{\partial \bar{x}^m}{\partial x^h} \bar{\Gamma}_m{}^j{}_p \bar{Y}_j \, d\bar{x}^p.$$

This is substituted in (3.24), in the first term of which the index $j$ is replaced by $m$. It is thus found that

$$dY_h - \Gamma_h{}^l{}_k Y_l \, dx^k = \frac{\partial \bar{x}^m}{\partial x^h} (d\bar{Y}_m - \bar{\Gamma}_m{}^j{}_p \bar{Y}_j \, d\bar{x}^p), \tag{3.25}$$

which shows that the *quantities defined by*

$$DY_h = dY_h - \Gamma_h{}^l{}_k Y_l \, dx^k \tag{3.26}$$

*are the components of a covariant vector*. We therefore regard (3.26) as the definition of the *absolute differential* of the covariant vector field $Y_h$.

In the next section we deal with the definition of absolute differentials of tensors of arbitrary type. However, we anticipate to some extent by stating

already at this stage that *the absolute differential of a scalar is defined to be its ordinary differential,* for as we have seen above, the latter is automatically a type (0, 0) tensor. Let us briefly examine the relevance of this definition with regard to the scalar product $\phi = X^h Y_h$ of two given co- and contravariant vector fields. By definition we have

$$D(X^h Y_h) = D\phi = d\phi = d(X^h Y_h) = (dX^h)Y_h + (X^h)\, dY_h.$$

On the right-hand side the differentials $dX^h$, $dY_h$ are replaced by means of the corresponding absolute differentials $DX^h$, $DY_h$ in accordance with (3.17) and (3.26), respectively:

$$D(X^h Y_h) = (DX^h - \Gamma_{l\,k}^{h} X^l\, dx^k)Y_h + X^h(DY_h + \Gamma_{h\,k}^{l} Y_l\, dx^k)$$
$$= (DX^h)Y_h + (X^h)\, DY_h, \tag{3.27}$$

where, in the second step, we have taken into account the fact that the two terms involving the $\Gamma_{l\,k}^{h}$ cancel (after an interchange of indices). This result suggests that the product rule of ordinary differentiation holds also for absolute differentials.

## 3.4 ABSOLUTE DIFFERENTIALS OF TENSOR FIELDS

We have seen that the ordinary differentials of co- or contravariant vector fields are not tensorial. This is true also for differentials of tensor fields of higher rank, as is immediately evident by differentiation of the transformation law (3.1) of a type $(r, s)$ tensor. However, if it is supposed that our differentiable manifold $X_n$ is endowed with a connection—as will be assumed here—it is a simple matter to construct the absolute differential of a tensor field of arbitrary type. This is done by differentiation of (3.1) and the subsequent elimination of all unwanted expressions by means of (3.15) and (3.22) in terms of the connection coefficients. Again, in order to avoid an unnecessarily complicated calculation, we illustrate this process for the case of a type (1, 1) tensor field.

Let us therefore consider a tensor field $T_h^p(x^q)$, whose transformation law is

$$\overline{T}_l^j = \frac{\partial \overline{x}^j}{\partial x^p}\frac{\partial x^h}{\partial \overline{x}^l} T_h^p. \tag{4.1}$$

Thus the differentials of $T_h^p$ relative to the two coordinate systems are related according to

$$d\overline{T}_l^j = \frac{\partial^2 \overline{x}^j}{\partial x^p\,\partial x^k}\frac{\partial x^h}{\partial \overline{x}^l} T_h^p\, dx^k + \frac{\partial \overline{x}^j}{\partial x^p}\frac{\partial^2 x^h}{\partial \overline{x}^l\,\partial \overline{x}^s} T_h^p\, d\overline{x}^s + \frac{\partial \overline{x}^j}{\partial x^p}\frac{\partial x^h}{\partial \overline{x}^l} dT_h^p. \tag{4.2}$$

Clearly the first two terms on the right-hand side of (4.2) are to be eliminated: we shall deal with each one of them in turn.

To the first term we apply (3.15), which gives

$$\frac{\partial^2 \overline{x}^j}{\partial x^p\,\partial x^k}\frac{\partial x^h}{\partial \overline{x}^l} T_h^p\, dx^k = \left(\frac{\partial \overline{x}^j}{\partial x^q}\Gamma_{p\,k}^{q} - \overline{\Gamma}_{r\,s}^{j}\frac{\partial \overline{x}^r}{\partial x^p}\frac{\partial \overline{x}^s}{\partial x^k}\right)\frac{\partial x^h}{\partial \overline{x}^l} T_h^p\, dx^k.$$

In the second expression on the right we invoke the transformation law (4.1), together with the inverse of (3.6); in both expressions various indices are relabeled. We thus obtain

$$\frac{\partial^2 \overline{x}^j}{\partial x^p\,\partial x^k}\frac{\partial x^h}{\partial \overline{x}^l} T_h^p\, dx^k = \frac{\partial \overline{x}^j}{\partial x^p}\frac{\partial x^h}{\partial \overline{x}^l}\Gamma_{m\,k}^{p} T_h^m\, dx^k - \overline{\Gamma}_{q\,s}^{j}\overline{T}_l^q\, d\overline{x}^s. \tag{4.3}$$

To the second term on the right-hand side of (4.2) we now apply (3.22) and proceed similarly:

$$\frac{\partial^2 x^h}{\partial \overline{x}^l\,\partial \overline{x}^s}\frac{\partial \overline{x}^j}{\partial x^p} T_h^p\, d\overline{x}^s = \frac{\partial x^h}{\partial \overline{x}^q}\overline{\Gamma}_{l\,s}^{q}\frac{\partial \overline{x}^j}{\partial x^p} T_h^p\, d\overline{x}^s - \frac{\partial x^r}{\partial \overline{x}^l}\frac{\partial x^k}{\partial \overline{x}^s}\Gamma_{r\,k}^{h}\frac{\partial \overline{x}^j}{\partial x^p} T_h^p\, d\overline{x}^s$$

$$= \overline{\Gamma}_{l\,s}^{q}\,\overline{T}_q^j\, d\overline{x}^s - \frac{\partial \overline{x}^j}{\partial x^p}\frac{\partial x^h}{\partial \overline{x}^l}\Gamma_{h\,k}^{m} T_m^p\, dx^k. \tag{4.4}$$

We can now substitute (4.3) and (4.4) in (4.2), placing all terms involving $\overline{T}$ and $\overline{\Gamma}$ on the left-hand side, which yields

$$d\overline{T}_l^j + \overline{\Gamma}_{q\,s}^{j}\overline{T}_l^q\, d\overline{x}^s - \overline{\Gamma}_{l\,s}^{q}\overline{T}_q^j\, d\overline{x}^s$$
$$= \frac{\partial \overline{x}^j}{\partial x^p}\frac{\partial x^h}{\partial \overline{x}^l}(dT_h^p + \Gamma_{m\,k}^{p} T_h^m\, dx^k - \Gamma_{h\,k}^{m} T_m^p\, dx^k). \tag{4.5}$$

Direct inspection of this result clearly indicates that the quantities defined by

$$DT_h^p = dT_h^p + (\Gamma_{m\,k}^{p} T_h^m - \Gamma_{h\,k}^{m} T_m^p)\, dx^k \tag{4.6}$$

represent the components of a type (1, 1) tensor field. Accordingly we regard (4.6) as the definition of the *absolute differential of a type* (1, 1) *tensor field.*

The same process, applied to the general transformation law (3.1), yields the following expression for the *absolute differential of a type* $(r, s)$ *tensor field*:

$$DT^{j_1\cdots j_r}{}_{l_1\cdots l_s} = dT^{j_1\cdots j_r}{}_{l_1\cdots l_s} + \sum_{\alpha=1}^{r}\Gamma_{m\,k}^{j_\alpha} T^{j_1\cdots j_{\alpha-1} m j_{\alpha+1}\cdots j_r}{}_{l_1\cdots l_s}\, dx^k$$

$$- \sum_{\beta=1}^{s}\Gamma_{l_\beta\,k}^{m} T^{j_1\cdots j_r}{}_{l_1\cdots l_{\beta-1} m l_{\beta+1}\cdots l_s}\, dx^k. \tag{4.7}$$

Clearly (3.17), (3.26), and (4.6) are special cases of this formula. For future reference we note also the following equally important particular cases:

first, for a type (0, 2) tensor field we have

$$DU_{lj} = dU_{lj} - \Gamma_l{}^m{}_k U_{mj}\, dx^k - \Gamma_j{}^m{}_k U_{lm}\, dx^k, \tag{4.8}$$

and second, for a type (2, 0) tensor field

$$DV^{lj} = dV^{lj} + \Gamma_m{}^l{}_k V^{mj}\, dx^k + \Gamma_m{}^j{}_k V^{lm}\, dx^k. \tag{4.9}$$

We are now in a position to state the *basic laws of absolute differentiation*:

1. The absolute differential of a type $(r, s)$ tensor field is a type $(r, s)$ tensor; in particular, the absolute differential of a scalar field is its ordinary differential (so that the absolute differential of a constant vanishes identically).

2. The absolute differential of the sum of two tensor fields of the same type is the sum of the absolute differentials of these fields.

3. The absolute differential of the product of any two tensor fields is given in terms of the absolute differentials of each of these fields by a rule which is formally identical with the product rule of ordinary differentiation.

Statement 1 is a direct consequence of our construction, while 2 follows directly from the linearity of the right-hand side of (4.7) in $T^{\cdots}\cdots$. The product rule 3, of which we have already encountered an example in (3.27), can also be derived in full generality directly from (4.7). Again, in order to avoid a profusion of indices, let us verify it for a special case, namely for the product of a type (2, 0) tensor field with a type (0, 1) field. Accordingly we shall consider the product

$$T^{lj}{}_h = V^{lj} Y_h. \tag{4.10}$$

From (4.7) we have

$$D(T^{lj}{}_h) = d(T^{lj}{}_h) + (\Gamma_m{}^l{}_k T^{mj}{}_h + \Gamma_m{}^j{}_k T^{lm}{}_h - \Gamma_h{}^m{}_k T^{lj}{}_m)\, dx^k, \tag{4.11}$$

in which we substitute from (4.10), which gives

$$\begin{aligned} D(V^{lj} Y_h) = {}&(dV^{lj}) Y_h + V^{lj}(dY_h) \\ &+ (\Gamma_m{}^l{}_k V^{mj} + \Gamma_m{}^j{}_k V^{lm})\, dx^k Y_h - V^{lj}(\Gamma_h{}^m{}_k Y_m\, dx^k). \end{aligned} \tag{4.12}$$

From (4.9) it is evident that the coefficient of $Y_h$ in this expression is simply $DV^{lj}$, while according to (3.26), the coefficient of $V^{lj}$ is $DY_h$. Thus (4.12) reduces to

$$D(V^{lj} Y_h) = (DV^{lj}) Y_h + V^{lj}(DY_h) \tag{4.13}$$

as required.

Before concluding this section we remark that the above definitions and rules of absolute differentiation are valid for *any* set of connection coefficients $\Gamma_h{}^l{}_k(x)$. *Our conclusions depend solely on the transformation law* (3.16) *of the connection coefficients.* For instance, if we are given another

connection, say $\tilde{\Gamma}_h{}^l{}_k$, satisfying the same transformation law, namely,

$$\bar{\tilde{\Gamma}}_m{}^j{}_p = \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p} \tilde{\Gamma}_h{}^l{}_k - \frac{\partial^2 \bar{x}^j}{\partial x^h\, \partial x^k} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p}, \tag{4.14}$$

we can use these alternative coefficients to construct absolute differentials of tensor fields as in (4.7): naturally the numerical values of the components of these absolute differentials will differ from those previously constructed, but we will again obtain tensors of the required type. This is also immediately evident from the fact that the right-hand side of (4.7) is linear in the connection coefficients, for it follows immediately by subtraction of (4.14) from (3.16) that *the differences of the respective coefficients of two connections constitute the components of a type* (1, 2) *tensor field*:

$$\bar{\tilde{\Gamma}}_m{}^j{}_p - \bar{\Gamma}_m{}^j{}_p = \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial x^h}{\partial \bar{x}^m} \frac{\partial x^k}{\partial \bar{x}^p} (\tilde{\Gamma}_h{}^l{}_k - \Gamma_h{}^l{}_k). \tag{4.15}$$

This conclusion entails further important consequences. Let us decompose a given affine connection $\Gamma_h{}^l{}_k$ into its symmetric and skew-symmetric parts according to the usual rule:

$$\Gamma_h{}^l{}_k = \tfrac{1}{2}(\Gamma_h{}^l{}_k + \Gamma_k{}^l{}_h) + \tfrac{1}{2}(\Gamma_h{}^l{}_k - \Gamma_k{}^l{}_h). \tag{4.16}$$

The connection is said to be *symmetric* if

$$\Gamma_h{}^l{}_k = \Gamma_k{}^l{}_h, \tag{4.17}$$

that is, if its skew-symmetric part vanishes. However, if $\Gamma_h{}^l{}_k$ represents a connection, so does $\Gamma_k{}^l{}_h$, as is immediately evident from (3.15) in view of the symmetry of $\partial^2 \bar{x}^j / \partial x^h\, \partial x^k$ in the indices $h$ and $k$. Thus, as a special case of (4.15) with $\tilde{\Gamma}_h{}^l{}_k = \Gamma_k{}^l{}_h$, we infer that

$$\Gamma_h{}^l{}_k - \Gamma_k{}^l{}_h \tag{4.18}$$

is a tensor. This is often referred to as the *torsion tensor* of the connection. Clearly, if (4.18) vanishes in some coordinate system, it will vanish in any other system, and accordingly *the symmetry condition* (4.17) *is independent of the choice of the coordinate system.*

## 3.5  PARTIAL COVARIANT DIFFERENTIATION

The relation between the differential $df$ of a function $f(x^k)$ of class $C^1$ and the partial derivatives of $f$ is given by $df = (\partial f / \partial x^k)\, dx^k$. This immediately raises the question as to the existence of a counterpart of this relation for absolute differentials: clearly it is necessary, for this purpose, to construct a

tensorial analogue of partial derivatives. Again, this is easily achieved by means of a given connection.

For the sake of simplicity we shall consider once more a contravariant vector field $X^h = X^h(x^p)$ of class $C^1$ on $X_n$, whose transformation law is given by

$$\overline{X}^j(\overline{x}^l) = \frac{\partial \overline{x}^j}{\partial x^h} X^h(x^p). \tag{5.1}$$

In the preceding section we saw that the partial derivatives of $X^h(x^p)$ are not tensorial; indeed, according to (3.5) these derivatives transform according to

$$\frac{\partial \overline{X}^j}{\partial \overline{x}^l} = \frac{\partial^2 \overline{x}^j}{\partial x^h \partial x^k} \frac{\partial x^k}{\partial \overline{x}^l} X^h + \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial x^k}{\partial \overline{x}^l} \frac{\partial X^h}{\partial x^k}. \tag{5.2}$$

Again we eliminate the first term on the right-hand side by means of (3.15). To this end, we multiply (3.15) by $(\partial x^k/\partial \overline{x}^l)X^h$, after which we apply (1.4) and (5.1) to the second term on the right:

$$\frac{\partial^2 \overline{x}^j}{\partial x^h \partial x^k} \frac{\partial x^k}{\partial \overline{x}^l} X^h = \frac{\partial \overline{x}^j}{\partial x^p} \Gamma_h{}^p{}_k \frac{\partial x^k}{\partial \overline{x}^l} X^h - \overline{\Gamma}_m{}^j{}_p \frac{\partial \overline{x}^m}{\partial x^h} \frac{\partial \overline{x}^p}{\partial x^k} \frac{\partial x^k}{\partial \overline{x}^l} X^h$$

$$= \frac{\partial \overline{x}^j}{\partial x^p} \frac{\partial x^k}{\partial \overline{x}^l} \Gamma_h{}^p{}_k X^h - \overline{\Gamma}_m{}^j{}_l \overline{X}^m. \tag{5.3}$$

This is substituted in (5.2), which yields

$$\frac{\partial \overline{X}^j}{\partial \overline{x}^l} + \overline{\Gamma}_m{}^j{}_l \overline{X}^m = \frac{\partial \overline{x}^j}{\partial x^p} \frac{\partial x^k}{\partial \overline{x}^l} \left( \frac{\partial X^p}{\partial x^k} + \Gamma_h{}^p{}_k X^h \right),$$

showing that the quantities defined by

$$X^p_{|k} = \frac{\partial X^p}{\partial x^k} + \Gamma_h{}^p{}_k X^h \tag{5.4}$$

constitute the components of a type (1, 1) tensor. We refer to (5.4) as the *covariant derivative of the contravariant vector field* $X^p(x^k)$. Since $dX^p = (\partial X^p/\partial x^k)\, dx^k$, it follows immediately from (5.4) and (3.17) that the absolute differential $DX^p$ is given by

$$DX^p = X^p_{|k}\, dx^k \tag{5.5}$$

for *any* displacement $dx^k$. This is the counterpart of the relation referred to in the opening sentence of this section which we have been seeking.

[Many texts use the following notations for (5.4): $X^p_{,k}$; $X^p_{;k}$; $\nabla_k X^p$; we adhere consistently to the notation given in (5.4).]

Clearly a similar process can be applied to a covariant vector field $Y_h(x^k)$, and it is found that the covariant derivative of this field may be defined as

$$Y_{h|k} = \frac{\partial Y_h}{\partial x^k} - \Gamma_h{}^m{}_k Y_m, \tag{5.6}$$

for which we again have

$$DY_h = Y_{h|k}\, dx^k$$

by virtue of (3.26).

*The (partial) covariant derivative of a type* (r, s) *tensor field* $T^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}$ *is defined to be*

$$T^{j_1 \cdots j_r}{}_{l_1 \cdots l_s|k} = \frac{\partial T^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}}{\partial x^k} + \sum_{\alpha=1}^{r} \Gamma_m{}^{j_\alpha}{}_k T^{j_1 \cdots j_{\alpha-1} m j_{\alpha+1} \cdots j_r}{}_{l_1 \cdots l_s}$$

$$- \sum_{\beta=1}^{s} \Gamma_{l_\beta}{}^m{}_k T^{j_1 \cdots j_r}{}_{l_1 \cdots l_{\beta-1} m l_{\beta+1} \cdots l_s}, \tag{5.7}$$

so that, according to (4.7), we have

$$T^{j_1 \cdots j_r}{}_{l_1 \cdots l_s|k}\, dx^k = DT^{j_1 \cdots j_r}{}_{l_1 \cdots l_s} \tag{5.8}$$

for any displacement $dx^k$.

The laws of partial covariant differentiation may be summarized as follows:

1.   The covariant derivative of a type (r, s) tensor field is a type (r, s + 1) tensor field; in particular the covariant derivative of a scalar is its ordinary derivative (so that the covariant derivative of a constant vanishes identically).

2.   The covariant derivative of the sum of two tensor fields of the same type is the sum of the covariant derivatives of these fields.

3.   The covariant derivative of the product of any two tensor fields is given in terms of the covariant derivatives of these fields by a rule which is formally identical with the product rule of partial differentiation in the usual sense.

Statement 1 is an immediate consequence of our construction, while 2 follows directly from the linearity of the right-hand side of (5.7) in $T^{\cdots}{}_{\cdots}$. The product rule 3 may be verified by direct calculation as in the case of absolute differentials. Thus operations with covariant derivatives are very similar to those of ordinary differentiation. There is, however, a certain aspect which entails a distinction of profound importance. This involves the order of repeated covariant differentiation, which will be treated in the next

section. Another feature which is also peculiar to covariant differentiation may be illustrated by the following important special case.

Suppose, for the moment, that our differentiable manifold $X_n$ is endowed with a type (0, 2) tensor field $a_{hk}(x^l)$ of class $C^1$ on $X_n$, it being assumed that this tensor field is symmetric:

$$a_{hk}(x^l) = a_{kh}(x^l). \qquad (5.9)$$

The transformation law satisfied by this field is

$$a_{hk}(x^l) = \bar{a}_{jp}(\bar{x}^m) \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k}. \qquad (5.10)$$

Let us differentiate this relation with respect to $x^l$, using the chain rule:

$$\frac{\partial a_{hk}}{\partial x^l} = \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} \frac{\partial \bar{x}^m}{\partial x^l} + \bar{a}_{jp} \frac{\partial^2 \bar{x}^j}{\partial x^l \partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} + \bar{a}_{jp} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial^2 \bar{x}^p}{\partial x^l \partial x^k}. \qquad (5.11)$$

We shall endeavor to solve this system for the second derivatives $\partial^2 \bar{x}^j / \partial x^h \partial x^k$. In order to do so, we permute the indices $h$, $k$, $l$ cyclically, which gives rise to two additional counterparts of (5.11):

$$\frac{\partial a_{kl}}{\partial x^h} = \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial \bar{x}^p}{\partial x^l} \frac{\partial \bar{x}^m}{\partial x^h} + \bar{a}_{jp} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^p}{\partial x^l} + \bar{a}_{jp} \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial^2 \bar{x}^p}{\partial x^h \partial x^l}, \qquad (5.12)$$

and

$$\frac{\partial a_{lh}}{\partial x^k} = \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial \bar{x}^p}{\partial x^h} \frac{\partial \bar{x}^m}{\partial x^k} + \bar{a}_{jp} \frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^l} \frac{\partial \bar{x}^p}{\partial x^h} + \bar{a}_{jp} \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial^2 \bar{x}^p}{\partial x^k \partial x^h}. \qquad (5.13)$$

Here it is to be noted that, by virtue of the symmetry assumption (5.9), the second term on the right-hand side of (5.11) is identical with the last term on the right-hand side of (5.12), for the former may be written as

$$\bar{a}_{jp} \frac{\partial \bar{x}^p}{\partial x^k} \frac{\partial^2 \bar{x}^j}{\partial x^l \partial x^h} = \bar{a}_{pj} \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial^2 \bar{x}^p}{\partial x^h \partial x^l} = \bar{a}_{jp} \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial^2 \bar{x}^p}{\partial x^h \partial x^l}.$$

Similarly, the third term on the right-hand side of (5.11) may be identified with the second term on the right-hand side of (5.13), while the remaining two terms in (5.12) and (5.13) involving the second derivatives of $\bar{x}^j$ and $\bar{x}^p$ are also respectively identical. Thus if we subtract (5.11) from the sum of (5.12) and (5.13), it is found that

$$\left( \frac{\partial a_{kl}}{\partial x^h} + \frac{\partial a_{lh}}{\partial x^k} - \frac{\partial a_{hk}}{\partial x^l} \right) = 2\bar{a}_{jp} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^p}{\partial x^l} + \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial \bar{x}^p}{\partial x^l} \frac{\partial \bar{x}^m}{\partial x^h}$$

$$+ \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial \bar{x}^p}{\partial x^h} \frac{\partial \bar{x}^m}{\partial x^k} - \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} \frac{\partial \bar{x}^m}{\partial x^l}.$$

After a suitable change of the indices $j$, $p$, $m$ in the second and third terms on the right we may write this relation in the form

$$\frac{1}{2}\left( \frac{\partial a_{kl}}{\partial x^h} + \frac{\partial a_{lh}}{\partial x^k} - \frac{\partial a_{hk}}{\partial x^l} \right) = \bar{a}_{jp} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \frac{\partial \bar{x}^p}{\partial x^l}$$

$$+ \frac{1}{2} \left( \frac{\partial \bar{a}_{pm}}{\partial \bar{x}^j} + \frac{\partial \bar{a}_{mj}}{\partial \bar{x}^p} - \frac{\partial \bar{a}_{jp}}{\partial \bar{x}^m} \right) \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \bar{x}^p}{\partial x^k} \frac{\partial \bar{x}^m}{\partial x^l}. \qquad (5.14)$$

This suggests that, in analogy with (2.5.14), we define the *Christoffel symbols of the first kind with respect to the symmetric tensor field* $a_{hk}(x^l)$ as follows:

$$\overset{(a)}{\Gamma}_{hlk} = \frac{1}{2}\left( \frac{\partial a_{kl}}{\partial x^h} + \frac{\partial a_{lh}}{\partial x^k} - \frac{\partial a_{hk}}{\partial x^l} \right), \qquad (5.15)$$

where it is to be noted that the superscript $(a)$ is intended to emphasize the fact that these quantities are defined by means of the given tensor field $a_{hk}(x^l)$. Thus, after changing a few indices, we can write (5.14) in the form

$$\overset{(a)}{\Gamma}_{hlk} = \overset{(a)}{\Gamma}_{qmp} \frac{\partial \bar{x}^q}{\partial x^h} \frac{\partial \bar{x}^m}{\partial x^l} \frac{\partial \bar{x}^p}{\partial x^k} + \bar{a}_{qp} \frac{\partial^2 \bar{x}^q}{\partial x^h \partial x^k} \frac{\partial \bar{x}^p}{\partial x^l}, \qquad (5.16)$$

from which it is immediately evident that the Christoffel symbols are not tensorial.

However, this relation is strongly reminiscent of the general transformation law (3.16) of a set of connection coefficients, and we now show how (5.16) may be used to construct a particular connection from the given tensor field $a_{hk}$, provided that this field is nonsingular in the sense that $a_{hk}$ possesses an inverse $a^{hl}$ everywhere:

$$a_{hk}a^{hl} = \delta_k^l. \qquad (5.17)$$

Since $a_{hk}$ is a type (0, 2) tensor (see Problem 3.23), its inverse $a^{hl}$ is a type (2, 0) tensor, whose transformation law is given by

$$a^{lj} = \bar{a}^{sr} \frac{\partial x^l}{\partial \bar{x}^s} \frac{\partial x^j}{\partial \bar{x}^r}. \qquad (5.18)$$

If this is multiplied by $\partial \bar{x}^m / \partial x^l$, the identity (1.4) being taken into account, it is found that

$$a^{lj} \frac{\partial \bar{x}^m}{\partial x^l} = \bar{a}^{sr} \delta_s^m \frac{\partial x^j}{\partial \bar{x}^r} = \bar{a}^{mr} \frac{\partial x^j}{\partial \bar{x}^r}. \qquad (5.19)$$

We now multiply (5.16) by $a^{lj}$, applying (5.19) twice on the right-hand side:

$$a^{lj}\overset{(a)}{\Gamma}_{hlk} = \overset{(a)}{\Gamma}_{qmp}\frac{\partial\bar{x}^q}{\partial x^h}\frac{\partial\bar{x}^p}{\partial x^k}\left(a^{lj}\frac{\partial\bar{x}^m}{\partial x^l}\right) + \bar{a}_{qp}\left(a^{lj}\frac{\partial\bar{x}^p}{\partial x^l}\right)\frac{\partial^2\bar{x}^q}{\partial x^h\,\partial x^k}$$

$$= \bar{a}^{mr}\overset{(a)}{\Gamma}_{qmp}\frac{\partial x^j}{\partial\bar{x}^r}\frac{\partial\bar{x}^q}{\partial x^h}\frac{\partial\bar{x}^p}{\partial x^k} + \bar{a}_{qp}\bar{a}^{pr}\frac{\partial^2\bar{x}^q}{\partial x^h\,\partial x^k}\frac{\partial x^j}{\partial\bar{x}^r}$$

$$= \left(\bar{a}^{mr}\overset{(a)}{\Gamma}_{qmp}\right)\frac{\partial x^j}{\partial\bar{x}^r}\frac{\partial\bar{x}^q}{\partial x^h}\frac{\partial\bar{x}^p}{\partial x^k} + \frac{\partial x^j}{\partial\bar{x}^r}\frac{\partial^2\bar{x}^r}{\partial x^h\,\partial x^k},$$

where, in the last step, we have used (5.17). In analogy with (2.5.19) we now define the *Christoffel symbols of the second kind with respect to the symmetric tensor field* $a_{hk}(x^l)$ by

$$\overset{(a)}{\Gamma}{}_h{}^j{}_k = a^{lj}\overset{(a)}{\Gamma}_{hlk}, \tag{5.20}$$

and it follows directly that the transformation law of these quantities is given by

$$\overset{(a)}{\bar\Gamma}{}_h{}^j{}_k = \frac{\partial x^j}{\partial\bar{x}^r}\left(\frac{\partial\bar{x}^q}{\partial x^h}\frac{\partial\bar{x}^p}{\partial x^k}\overset{(a)}{\Gamma}{}_q{}^r{}_p + \frac{\partial^2\bar{x}^r}{\partial x^h\,\partial x^k}\right), \tag{5.21}$$

or alternatively, if we multiply by $\partial\bar{x}^m/\partial x^j$, applying (1.4) in this process,

$$\overset{(a)}{\bar\Gamma}{}_q{}^m{}_p\frac{\partial\bar{x}^q}{\partial x^h}\frac{\partial\bar{x}^p}{\partial x^k} = \frac{\partial\bar{x}^m}{\partial x^j}\overset{(a)}{\Gamma}{}_h{}^j{}_k - \frac{\partial^2\bar{x}^m}{\partial x^h\,\partial x^k}. \tag{5.22}$$

However, this is identical with the general transformation law (3.15) satisfied by any set of connection coefficients. It therefore follows that *the Christoffel symbols of the second kind defined as in (5.20) with respect to any symmetric nonsingular type (0, 2) tensor field of class $C^1$ represent the components of an affine connection on $X_n$.*

Moreover, it is immediately obvious from the definition (5.15) that

$$\overset{(a)}{\Gamma}_{hlk} = \overset{(a)}{\Gamma}_{klh}, \tag{5.23}$$

and hence, by (5.20),

$$\overset{(a)}{\Gamma}{}_h{}^j{}_k = \overset{(a)}{\Gamma}{}_k{}^j{}_h. \tag{5.24}$$

The Christoffel symbols therefore define a *symmetric connection* in the sense of (4.17), and, as we have seen at the end of the last section, this symmetry property is independent of the choice of the coordinates, despite the non-tensorial nature of the connection coefficients.

Also, if we permute the indices $h$ and $l$ in (5.15), we have

$$\overset{(a)}{\Gamma}_{lhk} = \frac{1}{2}\left(\frac{\partial a_{kh}}{\partial x^l} + \frac{\partial a_{hl}}{\partial x^k} - \frac{\partial a_{lk}}{\partial x^h}\right),$$

and if this is added to (5.15), the symmetry property (5.9) being observed, we obtain the useful identity

$$\overset{(a)}{\Gamma}_{hlk} + \overset{(a)}{\Gamma}_{lhk} = \frac{\partial a_{hl}}{\partial x^k}. \tag{5.25}$$

Moreover the covariant derivative of the given tensor field $a_{hl}(x^k)$ with respect to the Christoffel symbols (5.20) is given by

$$a_{hl|k} = \frac{\partial a_{hl}}{\partial x^k} - a_{ml}\overset{(a)}{\Gamma}{}_h{}^m{}_k - a_{hm}\overset{(a)}{\Gamma}{}_l{}^m{}_k,$$

or, if we use the inverse of (5.20),

$$a_{hl|k} = \frac{\partial a_{hl}}{\partial x^k} - \overset{(a)}{\Gamma}_{hlk} - \overset{(a)}{\Gamma}_{lhk}.$$

It therefore follows from (5.25) that

$$a_{hl|k} = 0 \tag{5.26}$$

identically. It has thus been shown that *the covariant derivative of any symmetric nonsingular type (0, 2) tensor field vanishes identically whenever this covariant derivative is defined in terms of the connection whose coefficients are the Christoffel symbols defined with respect to the given tensor field.* Similarly

$$a^{hl}{}_{|k} = 0; \tag{5.27}$$

it should be remarked, however, that the identities (5.26), (5.27) are in general only valid for the special connection constructed here.

## 3.6  REPEATED COVARIANT DIFFERENTIATION

We have seen that the partial covariant derivative of a type $(r, s)$ tensor field gives rise to a type $(r, s + 1)$ tensor field. Clearly the latter may be differentiated once more, yielding a type $(r, s + 2)$ tensor field. For instance, given a class $C^2$ contravariant vector field $X^j(x^h)$, we may construct the covariant derivatives $X^j_{|h}$ and subsequently $X^j_{|h|k}$. This is analogous to forming the second partial derivatives $\partial^2\phi/\partial x^h\,\partial x^k$ of a scalar function $\phi(x^h)$ which is twice continuously differentiable in all its arguments. However, for such a

function the symmetry relation

$$\frac{\partial^2 \phi}{\partial x^h \partial x^k} = \frac{\partial^2 \phi}{\partial x^k \partial x^h} \tag{6.1}$$

holds identically, and the question now arises as to whether the order of repeated partial *covariant* differentiation of tensor fields is also immaterial.

It will be seen that the answer to this question is generally in the negative; indeed, this fact gives rise to a distinction between the processes of ordinary and covariant partial differentiation which is of paramount importance. We shall now consider this situation in some detail for the case of a contra-variant vector field whose components $X^j(x^h)$ are class $C^2$ functions of the coordinates. Again we have to assume that our differentiable manifold $X_n$ is endowed with a connection $\Gamma_l{}^j{}_h$, so that, according to (5.7), we can form the covariant derivatives

$$X^j_{|h} = \frac{\partial X^j}{\partial x^h} + \Gamma_l{}^j{}_h X^l, \tag{6.2}$$

and

$$X^j_{|h|k} = \frac{\partial}{\partial x^k}(X^j_{|h}) + \Gamma_m{}^j{}_k(X^m_{|h}) - \Gamma_h{}^l{}_k(X^j_{|l}). \tag{6.3}$$

In analogy with (6.1) we must now examine the symmetry properties—or rather, the lack thereof—of the tensor $X^j_{|h|k}$ in the indices $h$ and $k$. To this end we have to analyze (6.3) with some care. Thus we substitute (6.2) in the first two terms on the right-hand side of (6.3) and carry out the required differentiation of the first term. (It will be evident almost immediately why we leave the third term in (6.3) unchanged.) It is thus found that

$$X^j_{|h|k} = \frac{\partial}{\partial x^k}\left(\frac{\partial X^j}{\partial x^h} + \Gamma_l{}^j{}_h X^l\right) + \Gamma_m{}^j{}_k\left(\frac{\partial X^m}{\partial x^h} + \Gamma_l{}^m{}_h X^l\right) - \Gamma_h{}^l{}_k(X^j_{|l})$$

$$= \frac{\partial^2 X^j}{\partial x^k \partial x^h} + \frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} X^l + \Gamma_l{}^j{}_h \frac{\partial X^l}{\partial x^k} + \Gamma_m{}^j{}_k \frac{\partial X^m}{\partial x^h}$$

$$+ \Gamma_m{}^j{}_k \Gamma_l{}^m{}_h X^l - \Gamma_h{}^m{}_k X^j_{|m}. \tag{6.4}$$

A similar result is obtained when the indices $h$, $k$ are interchanged in this relation, namely

$$X^j_{|k|h} = \frac{\partial^2 X^j}{\partial x^h \partial x^k} + \frac{\partial \Gamma_l{}^j{}_k}{\partial x^h} X^l + \Gamma_l{}^j{}_k \frac{\partial X^l}{\partial x^h} + \Gamma_m{}^j{}_h \frac{\partial X^m}{\partial x^k} + \Gamma_m{}^j{}_h \Gamma_l{}^m{}_k X^l - \Gamma_k{}^l{}_h X^j_{|l}. \tag{6.5}$$

Equation (6.5) is now subtracted from (6.4). In the course of this process we note, firstly, that the ordinary partial derivatives $\partial^2 X^j/\partial x^k \partial x^h$ are symmetric

since $X^j(x^h) \in C^2$, and secondly, that a change of the indices $l$ and $m$ in the third and fourth terms, respectively, on the right-hand side of (6.5) gives rise to the expression

$$\Gamma_m{}^j{}_k \frac{\partial X^m}{\partial x^h} + \Gamma_l{}^j{}_h \frac{\partial X^l}{\partial x^k},$$

which is identical with the sum of the fourth and third terms on the right-hand side of (6.4). Accordingly it is found that

$$X^j_{|h|k} - X^j_{|k|h} = \left(\frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^j{}_k}{\partial x^h} + \Gamma_m{}^j{}_k \Gamma_l{}^m{}_h - \Gamma_m{}^j{}_h \Gamma_l{}^m{}_k\right) X^l - (\Gamma_h{}^l{}_k - \Gamma_k{}^l{}_h) X^j_{|l}. \tag{6.6}$$

We now recall that, according to (4.18), the coefficient of $X^j_{|l}$ in the last expression on the right-hand side is a tensor, the so-called torsion tensor, to be denoted as follows:

$$S_h{}^l{}_k = \Gamma_h{}^l{}_k - \Gamma_k{}^l{}_h. \tag{6.7}$$

A specific notation is also introduced for the coefficient of $X^l$ in (6.6), namely

$$K_l{}^j{}_{hk} = \frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^j{}_k}{\partial x^h} + \Gamma_m{}^j{}_k \Gamma_l{}^m{}_h - \Gamma_m{}^j{}_h \Gamma_l{}^m{}_k. \tag{6.8}$$

Thus if (6.7) and (6.8) are substituted in (6.6), we finally obtain

$$X^j_{|h|k} - X^j_{|k|h} = K_l{}^j{}_{hk} X^l - S_h{}^l{}_k X^j_{|l}. \tag{6.9}$$

This is the relation which we have been seeking; it should be scrutinized in some detail. It is known that the left-hand side is a type $(1, 2)$ tensor, while the same is true of the second term on the right-hand side. Accordingly the remaining term, namely, $K_l{}^j{}_{hk} X^l$, is a type $(1, 2)$ tensor, and since the vector field $X^l(x^h)$ is completely arbitrary, it is inferred that *the quantities defined by* (6.8) *constitute the components of a type* $(1, 3)$ *tensor, the so-called curvature tensor of* $X_n$. [The reason for this nomenclature will become evident presently. Also, the reader should verify the tensorial nature of (6.8) by direct calculation, using the transformation law of the connection coefficients.]      (3.16)

It will be seen that, in general, the curvature tensor does not vanish. Therefore, even for a symmetric connection, for which $S_h{}^l{}_k = 0$, one is compelled to conclude from (6.9) that one must always distinguish between the tensors $X^j_{|h|k}$ and $X^j_{|k|h}$. This state of affairs can be interpreted geometrically, as will be seen in the next section. Furthermore, similar conclusions may also be reached for other types of tensor fields; for instance,

for a class $C^2$ covariant vector field $Y_j(x^h)$ one can prove in a similar manner that

$$Y_{j|h|k} - Y_{j|k|h} = -K_{j\,hk}^{\ l}\, Y_l - S_{h\,k}^{\ l}\, Y_{j|l}. \tag{6.10}$$

Again, by direct calculation it is easily shown that for any class $C^2$ tensor field $T^{j_1\cdots j_r}{}_{l_1\cdots l_s}$ of type $(r, s)$ the general formula reads as follows:

$$T^{j_1\cdots j_r}{}_{l_1\cdots l_s|h|k} - T^{j_1\cdots j_r}{}_{l_1\cdots l_s|k|h}$$

$$= \sum_{\alpha=1}^{r} K_{m\,hk}^{\ j_\alpha}\, T^{j_1\cdots j_{\alpha-1} m j_{\alpha+1}\cdots j_r}{}_{l_1\cdots l_s}$$

$$- \sum_{\beta=1}^{s} K_{l_\beta\,hk}^{\ \ m}\, T^{j_1\cdots j_r}{}_{l_1\cdots l_{\beta-1} m l_{\beta+1}\cdots l_s} - S_{h\,k}^{\ m}\, T^{j_1\cdots j_r}{}_{l_1\cdots l_s|m}. \tag{6.11}$$

Particular cases of special importance are the following:

$$T^{jl}{}_{|h|k} - T^{jl}{}_{|k|h} = K_{m\,hk}^{\ j}\, T^{ml} + K_{m\,hk}^{\ l}\, T^{jm} - S_{h\,k}^{\ m}\, T^{jl}{}_{|m}, \tag{6.12}$$

and

$$T_{jl|h|k} - T_{jl|k|h} = -K_{j\,hk}^{\ m}\, T_{ml} - K_{l\,hk}^{\ m}\, T_{jm} - S_{h\,k}^{\ m}\, T_{jl|m}. \tag{6.13}$$

The relations (6.9)–(6.13) are often referred to as the *Ricci identities*.

## 3.7 PARALLEL VECTOR FIELDS

We shall now turn to some geometrical implications of the process of covariant differentiation. The focal point of our analysis will be the concept of parallelism on a differentiable manifold $X_n$ endowed with an affine connection (Levi-Civita [1]). Again we shall consider contravariant vectors in the first instance, and we shall endeavor to establish criteria—if possible—according to which two vectors defined at distinct points of $X_n$ can be regarded as being parallel to each other. Thus, since different points of $X_n$ are involved in this problem, it is evident that the connection coefficients $\Gamma_{h\,k}^{\ j}$ will enter into our analysis.

Let us first recall an analogous situation in an $n$-dimensional Euclidean space $E_n$ which was considered in Section 2.5, namely, the characterization of parallel vector fields in $E_n$ referred to curvilinear coordinates. It was seen that a field of parallel vectors of constant length can be described by the partial differential equations (2.5.22), namely

$$\frac{\partial X^j}{\partial x^k} + \{_{h\,k}^{\,j}\} X^h = 0, \tag{7.1}$$

in which the Christoffel symbols (2.5.19) are defined by means of (2.5.14) in terms of the symmetric tensor $g_{hk}$ associated with the curvilinear system, the determinant of this tensor being positive (see Problem 3.23). However, if one compares the definitions (2.5.19) and (2.5.14) with the relations (5.15) and (5.20), it becomes evident that the Christoffel symbols (2.5.19) are identical with the connection coefficients (5.20) constructed as in Section 3.5 for the case $a_{hk} = g_{hk}$; that is,

$$\{_{h\,k}^{\,j}\} = \overset{(g)}{\Gamma}_{h\,k}^{\ j}. \tag{7.2}$$

It follows that the partial differential equation (7.1) which characterizes the parallel vector field is simply of the form $X_{|k}^{j} = 0$, in which the partial covariant derivative is defined in terms of the special connection (7.2). In particular, for an arbitrary displacement $dx^k$ we then have by virtue of the identity (5.5)—which is valid for any connection—that

$$DX^j = X_{|k}^{j}\, dx^k = 0 \tag{7.3}$$

for a parallel vector field in $E_n$.

We shall allow ourselves to be guided by this analogy from Euclidean geometry in the course of our analysis on the manifold $X_n$. Let us consider a curve $C$ on $X_n$ defined parametrically by the equations $x^j = x^j(t)$, where $t$ denotes the parameter of the curve, it being supposed that the functions $x^j(t)$ are of class $C^1$. The curve $C$ passes through two fixed points $P_1$ and $P_2$ of $X_n$, these points corresponding respectively to parameter values $t_1$ and $t_2$ $(t_1 < t_2)$. The vector field

$$\dot{x}^j(t) = \frac{dx^j}{dt} \tag{7.4}$$

consists of tangents to $C$.

Now let us suppose that we are given a contravariant vector field $X^j(t)$ defined at each point of $C$ in the interval $t_1 \le t \le t_2$, it being assumed that the components $X^j(t)$ are continuously differentiable functions of the parameter $t$. [It should be noted that this does *not* imply that the $X^j$ can be represented as class $C^1$ functions $X^j = X^j(x^h)$ of the coordinates of $X_n$; this is possible if and only if the $X^j$ are defined over a finite $n$-dimensional region of $X_n$, whereas in the present context the $X^j$ are given merely along a one-dimensional locus.]

By analogy with (7.3) we shall now introduce the following definition:

*The vector field $X^j(t)$ is said to be parallel along the curve $C$ if it satisfies the differential equations*

$$DX^j = 0 \text{ along } C. \tag{7.5}$$

By virtue of the definition (3.17) of the absolute differential this condition is equivalent to

$$dX^j + \Gamma_{h\ k}^{\ j} X^h \, dx^k = 0, \tag{7.6}$$

or, in terms of derivatives with respect to the parameter $t$,

$$\frac{dX^j}{dt} + \Gamma_{h\ k}^{\ j} X^h \frac{dx^k}{dt} = 0. \tag{7.7}$$

In this formulation of the definition of parallelism the phrase "along the curve $C$" appears repeatedly. This is due to the fact that the explicit form of the differential equation (7.7) depends essentially on $C$; we shall see that this circumstance has implications of profound importance.

An equivalent description of this concept of parallelism may be presented as follows. Suppose that we are given a vector $X^j_{(1)}$ at the point $P_1$ of $C$, corresponding to the parameter value $t = t_1$. The vector $X^j_{(1)} + dX^j$ at the point $Q$ of $C$ corresponding to the parameter value $t + dt$ is said to result from the *parallel displacement of $X^j_{(1)}$ from the point $P$ to the point $Q$ if the corresponding increment $dX^j$ is given by* (7.6). The continuation of this process of parallel displacement of $X^j_{(1)}$ along $C$ from $P_1$ to $P_2$ is tantamount to the construction of a vector field $X^j(t)$ along $C$ which satisfies the differential equation (7.7) along $C$. This vector field is therefore parallel along $C$ according to the original definition. Clearly the process of parallel displacement of $X^j_{(1)}$ from $P_1$ to $P_2$ is equivalent to the integration of (7.7) along $C$; thus, if we denote the values of the components thus obtained at $P_2$ by $X^j_{(2)}$, we may write

$$X^j_{(2)} = X^j_{(1)} - \int_{C}^{\ \ t_2}_{t_1} \Gamma_{h\ k}^{\ j}(x^l) X^h(t) \dot{x}^k(t) \, dt. \tag{7.8}$$

Here it is to be clearly understood that the integration is to be performed along the given curve $C$, which implies, inter alia, that we must substitute in the integrand on the right-hand side of (7.8) for $x^l = x^l(t)$ and $\dot{x}^k = \dot{x}^k(t)$ as functions of $t$ as prescribed by the parametric equations defining $C$. From this, however, it follows that in general the values of $X^j_{(2)}$ at $P$ are crucially dependent on the given curve $C$: in fact, these values would be independent of $C$ if and only if

$$\Gamma_{h\ k}^{\ j} X^h \dot{x}^k \, dt \tag{7.9}$$

were an exact differential, and there is no reason to suppose that this is true. As an immediate consequence of this state of affairs we are confronted with the following phenomenon. Let us consider another class $C^1$ curve, denoted by $C^*$, which joins $P_1$ to $P_2$. Again starting with the vector $X^j_{(1)}$ at $P_1$, we proceed to displace this vector from $P_1$ to $P_2$ by *parallelism along $C^*$*. In the
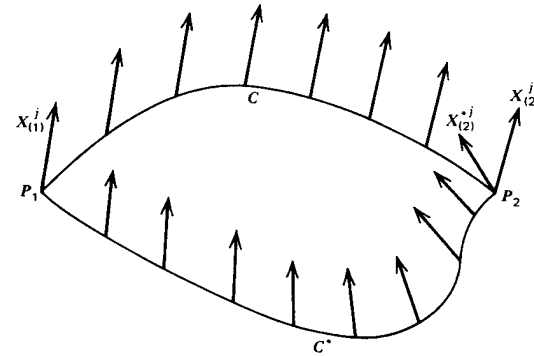
Fig. 6

relation corresponding to (7.8) the integrand assumes completely different values and therefore, in general, the vector $X^{*j}_{(2)}$ thus obtained at $P_2$ will differ from $X^j_{(2)}$ (see Figure 6):

$$X^{*j}_{(2)} \neq X^j_{(2)}. \tag{7.10}$$

This is in stark contrast with our notion of parallelism in Euclidean geometry. In the latter, given any two points $P_1$, $P_2$ and a vector $X^j_{(1)}$ at $P_1$, the vector at $P_2$ which is parallel to $X^j_{(1)}$ at $P_1$ is uniquely defined (except for its length), irrespective of any curve or curves joining $P_1$ to $P_2$. In the case of a general differentiable manifold endowed with a connection the corresponding situation is not nearly as simple. In fact, all that can be said in the present context is that the vector $X^j_{(2)}$ at the point $P_2$ of $X_n$ is obtained from the vector $X^j_{(1)}$ at the point $P_1$ of $X_n$ by parallel displacement *along the curve $C$*; similarly $X^{*j}_{(2)}$ is obtained by parallel displacement *along the curve $C^*$*. Thus in general *there is no absolute parallelism*; that is, there is no concept of parallelism which does not depend on the choice of some curve joining the points $P_1$ and $P_2$ at which the vectors $X^j_{(1)}$ and $X^j_{(2)}$ are to be compared.

It is only for points which are infinitesimally close to each other that this difficulty does not arise. For, given any contravariant vector $X^j$ at a point $P$ with coordinates $x^k$, a unique vector $X^j + dX^j$ at a neighboring point $Q$ with coordinates $(x^k + dx^k)$ may be defined by requiring that $dX^j$ be given by

$$dX^j = -\Gamma_{h\ k}^{\ j} X^h \, dx^k, \tag{7.11}$$

where $\Gamma_{h\ k}^{\ j}$ is to be evaluated at $P$, while $dx^k$ refers to the displacement $PQ$. Because of (7.6) it is consistent with our definition of parallelism to say that

the vector $X^j + dX^j$ thus constructed at $Q$ is *parallel* to the vector $X^j$ at $P$. Accordingly we always have what may be loosely called a form of *local parallelism*. In particular, since $X^j$ is an element of the tangent space $T_n(P)$ of $X_n$ at $P$, while $X^j + dX^j$ is an element of $T_n(Q)$, this local parallelism provides us with a method of relating elements of distinct tangent spaces belonging to neighboring points of $X_n$. In fact, by means of our affine connection $\Gamma_h{}^j{}_k$ a unique mapping of $T_n(P)$ onto the tangent space $T_n(Q)$ of a neighboring point is defined in this manner.

The phenomenon illustrated by the inequality (7.10) is referred to as the *nonintegrability of the parallel displacement*; this nomenclature is motivated by the fact that the integral on the right-hand side of (7.8) depends on the choice of the curve $C$.

The particular example of parallel vector fields in a Euclidean space indicates, however, that under certain special circumstances an integrable parallel displacement may be possible, and we shall now investigate the conditions which must be satisfied in order that this be feasible. Let us suppose, for the moment, that our $X_n$ is such as to permit the introduction of affine connection coefficients $\Gamma_h{}^j{}_k$ of class $C^1$ which are such that the parallel displacement of a given contravariant vector $X^j_{(1)}$ at $P_1$ leads to a *unique* vector $X^j$ at *any* other point $P$ of $X_n$; that is, the same vector $X^j$ is obtained at $P$ for all class $C^1$ curves joining $P_1$ to $P$. The coordinates of $P$ are denoted by $x^k$, and the vector $X^j$ thus constructed at $P$ will naturally also depend on these coordinates. Our construction therefore gives rise to a contravariant vector field $X^j = X^j(x^k)$ over a finite $n$-dimensional region $G$ of $X_n$ containing $P_1$. By construction, this field is of class $C^2$, so that its covariant derivative $X^j_{|k}$ is defined on $G$. Thus according to (5.5), we may write $DX^j = X^j_{|h} dx^h$ for any displacement $dx^h$, and, since our construction implies by virtue of (7.5) that $DX^j = 0$ for arbitrary values of $dx^h$, it follows that our parallel vector field satisfies the system of $n^2$ first-order partial differential equations

$$X^j_{|h} = 0 \tag{7.12}$$

on the region $G$. Written out in full, this system is

$$\frac{\partial X^j}{\partial x^h} = -\Gamma_l{}^j{}_h X^l. \tag{7.13}$$

As a result of our assumption that a class $C^2$ vector field $X^j$ exist such that (7.13) is satisfied, this implies that the partial derivatives of the left-hand side with respect to $x^k$ be symmetric in the indices $h$ and $k$. We therefore differentiate (7.13) with respect to $x^k$, which gives

$$\frac{\partial^2 X^j}{\partial x^k \, \partial x^h} = -\frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} X^l - \Gamma_l{}^j{}_h \frac{\partial X^l}{\partial x^k}.$$

In the second term on the right-hand side we must replace $\partial X^l / \partial x^k$ in accordance with (7.13), so that we obtain

$$\frac{\partial^2 X^j}{\partial x^k \, \partial x^h} = -\frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} X^l + \Gamma_l{}^j{}_h \Gamma_m{}^l{}_k X^m,$$

or

$$\frac{\partial^2 X^j}{\partial x^k \, \partial x^h} = \left( -\frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} + \Gamma_m{}^j{}_h \Gamma_l{}^m{}_k \right) X^l. \tag{7.14}$$

A similar relation is obtained by an interchange of the indices $h$ and $k$, from which (7.14) is subtracted, the postulated symmetry of the left-hand sides being taken into account. It is thus found that

$$\left( \frac{\partial \Gamma_l{}^j{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^j{}_k}{\partial x^h} + \Gamma_m{}^j{}_k \Gamma_l{}^m{}_h - \Gamma_m{}^j{}_h \Gamma_l{}^m{}_k \right) X^l = 0.$$

If we now invoke the definition (6.8) of the curvature tensor, we see that this is equivalent to

$$K_l{}^j{}_{hk} X^l = 0. \tag{7.15}$$

*This relation represents a necessary condition which must be satisfied by the vector field $X^l(x^k)$ if it is to be parallel in the sense of* (7.12)[†].

It is easily seen that this condition cannot be satisfied in general. Given an affine connection on $X_n$, the curvature tensor is uniquely defined by (6.8) at each point of $X_n$: the values of its components are therefore specified irrespective of the vector field $X^l$. Thus at the initial point $P_1$, for an arbitrary initial contravariant vector $X^l_{(1)}$, the left-hand side of (7.15) will not generally vanish, and by continuity it follows that (7.15) cannot be satisfied on a finite $n$-dimensional region containing $P_1$.

It should be remarked, however, that there are some exceptions. It is sometimes possible to construct certain special connections on a manifold for which the curvature tensor possesses particular properties which are such as to admit the existence of one or more special vector fields $X^j$ such that (7.15) is satisfied. Under these very special circumstances certain parallel vector fields may exist, but in this case, it is also necessary to investigate the integrability conditions of (7.15). Because of the specialized nature of these affine connections and vector fields we do not pursue this matter here.

A special case of more immediate interest is represented by a differentiable manifold $X_n$ on which one may construct a connection for which the curvature tensor vanishes identically. This phenomenon is independent of the

---

[†] Clearly (7.15) can also be derived directly by substitution of (7.12) in (6.9); however, the above derivation affords a better insight into the structure of the curvature tensor.

choice of the coordinate system on $X_n$, and under these circumstances $X_n$ is called a *flat* space. For instance, *the Euclidean space $E_n$ is flat in this sense.* This is due to the fact that $E_n$ admits a rectangular coordinate system in which the components of the tensor $g_{hj}$ are constants, so that, by (2.5.14) and (2.5.19), the Christoffel symbols (7.2) vanish everywhere in these coordinates, and according to (6.8) this entails that the components of the curvature tensor are identically zero.

However, if

$$K_l{}^j{}_{hk} = 0 \tag{7.16}$$

on $X_n$, it is obvious that (7.15) is satisfied automatically, and no further integrability conditions have to be considered. Under these circumstances the condition (7.15) is not only necessary but also sufficient for the existence of parallel vector fields, because (7.8) merely depends on the integral of an exact differential. This is easily verified by a calculation very similar to that giving rise to (7.14). *The condition (7.16) is therefore sufficient for the existence of parallel vector fields on $X_n$.* In fact, given an arbitrary initial contravariant vector $X^j_{(1)}$ at an arbitrary initial point $P_1$ of $X_n$, one can, if (7.16) is satisfied, construct a vector field $X^j(x^h)$ on a finite region containing $P_1$, each vector of which is parallel to $X^j_{(1)}$ at $P_1$ in the sense of (7.12).

As an example of the general concept of parallelism on a differentiable manifold $X_n$ let us consider the field of tangent vectors $\dot{x}^j(t)$ of a class $C^1$ curve $C$. The absolute differential of this field is given by

$$D\dot{x}^j = d\dot{x}^j + \Gamma_h{}^j{}_k \dot{x}^h \, dx^k,$$

and in accordance with our definition (7.5) the tangent vectors are parallel along $C$ if $D\dot{x}^j = 0$ along $C$, that is, if the functions $x^j = x^j(t)$ satisfy the system of $n$ second-order ordinary differential equations

$$\frac{d^2 x^j}{dt^2} + \Gamma_h{}^j{}_k \frac{dx^h}{dt} \frac{dx^k}{dt} = 0. \tag{7.17}$$

For a curve of this kind the tangent vectors are therefore generated by parallel displacement: this represents an obvious generalization of the corresponding property of the straight lines of Euclidean geometry. [Alternatively, a comparison of (7.17) with the differential equations (2.5.25) of straight lines in $E_n$ clearly points to the fact that (7.17) and (2.5.25) have formally the same structure.]

On a differentiable manifold $X_n$ endowed with a connection $\Gamma_h{}^j{}_k$ the curves satisfying the differential equations (7.17) play a special role; they are often referred to as *autoparallel* curves or *paths* of $X_n$. For this reason an affinely connected manifold $X_n$ is sometimes called a *space of paths* (and, in the past,

a *non-Riemannian* space) (see Eisenhart [3], Thomas [1]). The concept of paths on a manifold endowed with a connection should be regarded as a natural generalization of the notion of straight lines of Euclidean geometry.

The parameter $t$ which appears in (7.17) is always regarded as an invariant under coordinate transformations. However, it should be remarked that if $t$ is replaced by another parameter $\tau$ as a result of a class $C^2$ parameter transformation $t = t(\tau)$, where $\tau$ is again an invariant, the structure of (7.17) will be affected whenever $t(\tau)$ is a nonlinear function of $\tau$. Thus the characteristic structure (7.17) of the differential equations of the paths is dependent on the choice of the parameter $t$. This phenomenon is reflected also in the equation (2.5.25) of the straight lines of Euclidean geometry, in which the parameter $s$ is the arc length.

### 3.8 PROPERTIES OF THE CURVATURE TENSOR

In this section we derive some important algebraic and analytical properties of the curvature tensor as defined by (6.8). Although this definition appears to be somewhat complicated analytically, it gives rise directly to simple identities which are satisfied by the curvature tensor, and, when calculations involving the latter are performed, one usually uses these properties instead of the expression (6.8).

The first identity expresses the skew-symmetry of the curvature tensor in its last pair of indices:

$$K_l{}^j{}_{hk} = -K_l{}^j{}_{kh}, \tag{8.1}$$

which is immediately obvious from (6.8).

The second identity is obtained when one permutes the indices $l$, $h$, $k$ in (6.8) cyclically, thus obtaining two similar relations. These are added to (6.8), and after collecting terms in the manner indicated below, one obtains

$$K_l{}^j{}_{hk} + K_h{}^j{}_{kl} + K_k{}^j{}_{lh}$$

$$= \frac{\partial}{\partial x^k}(\Gamma_l{}^j{}_h - \Gamma_h{}^j{}_l) + \frac{\partial}{\partial x^l}(\Gamma_h{}^j{}_k - \Gamma_k{}^j{}_h) + \frac{\partial}{\partial x^h}(\Gamma_k{}^j{}_l - \Gamma_l{}^j{}_k)$$

$$+ \Gamma_m{}^j{}_k(\Gamma_l{}^m{}_h - \Gamma_h{}^m{}_l) + \Gamma_m{}^j{}_h(\Gamma_k{}^m{}_l - \Gamma_l{}^m{}_k) + \Gamma_m{}^j{}_l(\Gamma_h{}^m{}_k - \Gamma_k{}^m{}_h). \tag{8.2}$$

Here it is to be noted that the skew-symmetric parts of $\Gamma_h{}^j{}_k$ appear in a definite pattern. Accordingly we introduce the torsion tensor (6.7), so that (8.2) can be written in the form

$$K_l{}^j{}_{hk} + K_h{}^j{}_{kl} + K_k{}^j{}_{lh}$$

$$= \frac{\partial S_l{}^j{}_h}{\partial x^k} + \frac{\partial S_h{}^j{}_k}{\partial x^l} + \frac{\partial S_k{}^j{}_l}{\partial x^h} + \Gamma_m{}^j{}_k S_l{}^m{}_h + \Gamma_m{}^j{}_h S_k{}^m{}_l + \Gamma_m{}^j{}_l S_h{}^m{}_k. \tag{8.3}$$

The left-hand side of this relation is obviously tensorial, while the individual terms on the right-hand side are nontensorial. This suggests immediately that we must replace the partial derivatives of the torsion tensor by its covariant derivatives. We therefore recall that, by definition,

$$S_{l\ h|k}^{\ j} = \frac{\partial S_l^{\ j}_{\ h}}{\partial x^k} + \Gamma_m^{\ j}_{\ k}S_l^{\ m}_{\ h} - \Gamma_l^{\ m}_{\ k}S_m^{\ j}_{\ h} - \Gamma_h^{\ m}_{\ k}S_l^{\ j}_{\ m}. \tag{8.4}$$

Again the indices $l$, $h$, $k$ are permuted cyclically, and the relations thus obtained are added to (8.4). This yields

$$S_{l\ h|k}^{\ j} + S_{h\ k|l}^{\ j} + S_{k\ l|h}^{\ j}$$
$$= \frac{\partial S_l^{\ j}_{\ h}}{\partial x^k} + \frac{\partial S_h^{\ j}_{\ k}}{\partial x^l} + \frac{\partial S_k^{\ j}_{\ l}}{\partial x^h} + \Gamma_m^{\ j}_{\ k}S_l^{\ m}_{\ h} + \Gamma_m^{\ j}_{\ h}S_k^{\ m}_{\ l} + \Gamma_m^{\ j}_{\ l}S_h^{\ m}_{\ k} - S_l^{\ j}_{\ m}\Gamma_h^{\ m}_{\ k}$$
$$- S_m^{\ j}_{\ l}\Gamma_k^{\ m}_{\ h} - S_m^{\ j}_{\ h}\Gamma_l^{\ m}_{\ k} - S_h^{\ j}_{\ m}\Gamma_k^{\ m}_{\ l} - S_m^{\ j}_{\ k}\Gamma_h^{\ m}_{\ l} - S_k^{\ j}_{\ m}\Gamma_l^{\ m}_{\ h}. \tag{8.5}$$

The sum of the first six terms on the right-hand side of (8.5) is identical with the right-hand side of (8.3). Also, because of the skew-symmetry of the torsion tensor in its subscripts, the last six terms in (8.5) can be combined in pairs as follows:

$$-S_l^{\ j}_{\ m}(\Gamma_h^{\ m}_{\ k} - \Gamma_k^{\ m}_{\ h}) - S_h^{\ j}_{\ m}(\Gamma_k^{\ m}_{\ l} - \Gamma_l^{\ m}_{\ k}) - S_k^{\ j}_{\ m}(\Gamma_l^{\ m}_{\ h} - \Gamma_h^{\ m}_{\ l})$$
$$= -S_l^{\ j}_{\ m}S_h^{\ m}_{\ k} - S_h^{\ j}_{\ m}S_k^{\ m}_{\ l} - S_k^{\ j}_{\ m}S_l^{\ m}_{\ h}.$$

Thus when (8.5) is substituted in (8.3), it is found that

$$K_l^{\ j}_{\ hk} + K_h^{\ j}_{\ kl} + K_k^{\ j}_{\ lh} = S_{l\ h|k}^{\ j} + S_{h\ k|l}^{\ j} + S_{k\ l|h}^{\ j}$$
$$+ S_l^{\ j}_{\ m}S_h^{\ m}_{\ k} + S_h^{\ j}_{\ m}S_k^{\ m}_{\ l} + S_k^{\ j}_{\ m}S_l^{\ m}_{\ h}. \tag{8.6}$$

This is the identity which we have been seeking. In particular, for a *symmetric connection* ($S_l^{\ j}_{\ m} = 0$), one has

$$K_l^{\ j}_{\ hk} + K_h^{\ j}_{\ kl} + K_k^{\ j}_{\ lh} = 0, \tag{8.7}$$

which is already directly evident from (8.2).

The third identity which we shall require involves the covariant derivatives of the curvature tensor. In order to obtain this result we shall proceed somewhat indirectly as follows. Given some class $C^3$ covariant vector field $Y_j(x^k)$, let us differentiate (6.10) covariantly with respect to $x^p$:

$$Y_{j|h|k|p} - Y_{j|k|h|p} = -Y_l K_{j\ hk|p}^{\ l} - Y_{l|p}K_{j\ hk}^{\ l} - Y_{j|m}S_h^{\ m}_{\ k|p} - Y_{j|m|p}S_h^{\ m}_{\ k}. \tag{8.8}$$

Again a cyclic permutation of the indices $h$, $k$, $p$ is performed, and the resulting two equations are added to (8.8). In the relation thus obtained all terms are

collected as indicated, which yields the following result:

$$(Y_{j|h|k|p} - Y_{j|h|p|k}) + (Y_{j|k|p|h} - Y_{j|k|h|p}) + (Y_{j|p|h|k} - Y_{j|p|k|h})$$
$$= -Y_l(K_{j\ hk|p}^{\ l} + K_{j\ kp|h}^{\ l} + K_{j\ ph|k}^{\ l}) - K_{j\ hk}^{\ l}Y_{l|p} - K_{j\ kp}^{\ l}Y_{l|h} - K_{j\ ph}^{\ l}Y_{l|k}$$
$$- Y_{j|m}(S_h^{\ m}_{\ k|p} + S_k^{\ m}_{\ p|h} + S_p^{\ m}_{\ h|k})$$
$$- Y_{j|m|p}S_h^{\ m}_{\ k} - Y_{j|m|h}S_k^{\ m}_{\ p} - Y_{j|m|k}S_p^{\ m}_{\ h}. \tag{8.9}$$

To each of the three expressions on the left-hand side we apply the formula (6.13); accordingly this side becomes

$$-K_{j\ kp}^{\ l}Y_{l|h} - K_{h\ kp}^{\ l}Y_{j|l} - S_k^{\ m}_{\ p}Y_{j|h|m} - K_{j\ ph}^{\ l}Y_{l|k} - K_{k\ ph}^{\ l}Y_{j|l}$$
$$- S_p^{\ m}_{\ h}Y_{j|k|m} - K_{j\ hk}^{\ l}Y_{l|p} - K_{p\ hk}^{\ l}Y_{j|l} - S_h^{\ m}_{\ k}Y_{j|p|m}.$$

When this is substituted in (8.9), we see that the seventh, first, and fourth terms on the left cancel with the fourth, fifth, and sixth terms, respectively, on the right. Upon changing signs on each side, we obtain

$$Y_{j|l}(K_{h\ kp}^{\ l} + K_{k\ ph}^{\ l} + K_{p\ hk}^{\ l}) = Y_l(K_{j\ hk|p}^{\ l} + K_{j\ kp|h}^{\ l} + K_{j\ ph|k}^{\ l})$$
$$+ Y_{j|m}(S_h^{\ m}_{\ k|p} + S_k^{\ m}_{\ p|h} + S_p^{\ m}_{\ h|k})$$
$$+ S_h^{\ m}_{\ k}(Y_{j|m|p} - Y_{j|p|m}) + S_k^{\ m}_{\ p}(Y_{j|m|h} - Y_{j|h|m})$$
$$+ S_p^{\ m}_{\ h}(Y_{j|m|k} - Y_{j|k|m}). \tag{8.10}$$

To the last three expressions on the right-hand side we apply the identity (6.10), so that the sum of these expressions becomes

$$-S_h^{\ m}_{\ k}(K_{j\ mp}^{\ l}Y_l + S_m^{\ l}_{\ p}Y_{j|l}) - S_k^{\ m}_{\ p}(K_{j\ mh}^{\ l}Y_l + S_m^{\ l}_{\ h}Y_{j|l})$$
$$- S_p^{\ m}_{\ h}(K_{j\ mk}^{\ l}Y_l + S_m^{\ l}_{\ k}Y_{j|l})$$
$$= -Y_l(S_h^{\ m}_{\ k}K_{j\ mp}^{\ l} + S_k^{\ m}_{\ p}K_{j\ mh}^{\ l} + S_p^{\ m}_{\ h}K_{j\ mk}^{\ l})$$
$$- Y_{j|l}(S_h^{\ m}_{\ k}S_m^{\ l}_{\ p} + S_k^{\ m}_{\ p}S_m^{\ l}_{\ h} + S_p^{\ m}_{\ h}S_m^{\ l}_{\ k}).$$

This is substituted in (8.10), all terms containing $Y_{j|l}$ being collected on the left, the skew-symmetry of the torsion tensor being taken into account:

$$Y_{j|l}\{K_{h\ kp}^{\ l} + K_{k\ ph}^{\ l} + K_{p\ hk}^{\ l}$$
$$- S_h^{\ l}_{\ k|p} - S_k^{\ l}_{\ p|h} - S_p^{\ l}_{\ h|k} - S_h^{\ l}_{\ m}S_k^{\ m}_{\ p} - S_k^{\ l}_{\ m}S_p^{\ m}_{\ h} - S_p^{\ l}_{\ m}S_h^{\ m}_{\ k}\}$$
$$= Y_l\{K_{j\ hk|p}^{\ l} + K_{j\ kp|h}^{\ l} + K_{j\ ph|k}^{\ l} - S_h^{\ m}_{\ k}K_{j\ mp}^{\ l} - S_k^{\ m}_{\ p}K_{j\ mh}^{\ l} - S_p^{\ m}_{\ h}K_{j\ mk}^{\ l}\}.$$

However, by virtue of the identity (8.6) the coefficient of $Y_{j|l}$ on the left-hand side vanishes identically. Thus, since $Y_l$ is an entirely arbitrary vector field, the coefficient of $Y_l$ on the right must vanish also, so that

$$K_{j\ hk|p}^{\ l} + K_{j\ kp|h}^{\ l} + K_{j\ ph|k}^{\ l} = S_h^{\ m}_{\ k}K_{j\ mp}^{\ l} + S_k^{\ m}_{\ p}K_{j\ mh}^{\ l} + S_p^{\ m}_{\ h}K_{j\ mk}^{\ l}. \tag{8.11}$$

This is the so-called *Bianchi identity*: the importance of this identity in all applications of the tensor calculus cannot be overestimated. Again, in the

case of a *symmetric connection* this identity reduces to

$$K_{j\ hk|p}^{\ l} + K_{j\ kp|h}^{\ l} + K_{j\ ph|k}^{\ l} = 0. \tag{8.12}$$

Various new tensors can be obtained from the curvature tensor (6.8) by contraction. If we contract over the indices $j$ and $k$ we obtain a type $(0, 2)$ tensor, which is called the *Ricci tensor* and which is denoted as indicated:

$$K_{lh} = K_{l\ hj}^{\ j} = \frac{\partial \Gamma_{l\ h}^{\ j}}{\partial x^j} - \frac{\partial \Gamma_{l\ j}^{\ j}}{\partial x^h} + \Gamma_{m\ j}^{\ j}\Gamma_{l\ h}^{\ m} - \Gamma_{m\ h}^{\ j}\Gamma_{l\ j}^{\ m}. \tag{8.13}$$

In passing we note that if we contract over $j$ and $h$ in (6.8), no further new tensor is obtained since an application of (8.1) simply yields

$$K_{l\ jk}^{\ j} = -K_{l\ kj}^{\ j} = -K_{lk}. \tag{8.14}$$

Similarly, if we contract over $l$ and $j$ in (6.8) a tensor is obtained which, as we shall see, is also closely related to the Ricci tensor. First, from (6.8) we have

$$K_{l\ hk}^{\ l} = \frac{\partial \Gamma_{l\ h}^{\ l}}{\partial x^k} - \frac{\partial \Gamma_{l\ k}^{\ l}}{\partial x^h} + \Gamma_{m\ k}^{\ l}\Gamma_{l\ h}^{\ m} - \Gamma_{m\ h}^{\ l}\Gamma_{l\ k}^{\ m}.$$

Clearly, if we interchange the indices $l, m$ in the fourth term on the right-hand side, we see that it cancels with the third term, so that

$$K_{l\ hk}^{\ l} = \frac{\partial \Gamma_{l\ h}^{\ l}}{\partial x^k} - \frac{\partial \Gamma_{l\ k}^{\ l}}{\partial x^h}, \tag{8.15}$$

which is also a type $(0, 2)$ tensor. Second, if we contract over $j$ and $k$ in (8.6), noting (8.13) and (8.14), we obtain

$$K_{lh} - K_{hl} + K_{j\ lh}^{\ j} = S_{l\ h|j}^{\ j} + S_{h\ j|l}^{\ j} + S_{j\ l|h}^{\ j} + S_{l\ m}^{\ j}S_{h\ j}^{\ m} + S_{h\ m}^{\ j}S_{j\ l}^{\ m} + S_{j\ m}^{\ j}S_{l\ h}^{\ m}. \tag{8.16}$$

This suggests that we define the *torsion vector*

$$S_l = S_{j\ l}^{\ j} = -S_{l\ j}^{\ j}. \tag{8.17}$$

Also, the fifth term on the right-hand side of (8.16) can be written in the form

$$S_{h\ m}^{\ j}S_{j\ l}^{\ m} = S_{h\ j}^{\ m}S_{m\ l}^{\ j} = -S_{l\ m}^{\ j}S_{h\ j}^{\ m},$$

so that (8.16) reduces to

$$K_{j\ lh}^{\ j} = K_{hl} - K_{lh} + S_{l\ h|j}^{\ j} + S_{l|h} - S_{h|l} + S_jS_{l\ h}^{\ j}, \tag{8.18}$$

which is the required formula.

Thus for a symmetric connection ($S_{l\ h}^{\ j} = 0$), the tensor $\frac{1}{2}K_{j\ lh}^{\ j}$ is merely the skew-symmetric part of the Ricci tensor $K_{hl}$. In general, even in the case of a symmetric connection, the tensor (8.15) does not vanish. However, if there

exists a class $C^2$ function $f(x^h)$ of the coordinates such that

$$\Gamma_{j\ h}^{\ j} = \frac{\partial f}{\partial x^h}, \tag{8.19}$$

that is, if $\Gamma_{j\ h}^{\ j}$ is a gradient, then one has

$$\frac{\partial \Gamma_{j\ h}^{\ j}}{\partial x^l} = \frac{\partial \Gamma_{j\ l}^{\ j}}{\partial x^h}, \tag{8.20}$$

so that, by (8.15),

$$K_{j\ hl}^{\ j} = 0 \tag{8.21}$$

under these circumstances. It follows from (8.18) that *if a symmetric connection is such that its contracted coefficients $\Gamma_{j\ h}^{\ j}$ are the components of a gradient, the corresponding Ricci tensor is symmetric.*

In passing we note the following phenomenon. As a result of the transformation law (3.16) of the connection coefficients, the $\Gamma_{j\ h}^{\ j}$ do not form the components of a covariant vector field. It follows that the function $f(x^h)$ whose derivatives appear in (8.19) is not a scalar.

## PROBLEMS

Unless otherwise stated, all quantities are the components of tensors of the type indicated by the position of the indices.

3.1   Prove that   (a) $\delta_i^h \delta_k^j \delta_j^i = \delta_k^h$,

               (b) $\delta_j^i \delta_k^j \delta_i^k = n$.

3.2   Let $A^{hj}$ be a type $(2, 0)$ tensor, while $B_h$, $C_j$ denote type $(0, 1)$ tensors. Show that

$$A^{hj}B_hC_j$$

is a scalar.

3.3   Prove that none of the following are the components of tensors

$$A_{ij} + B^{ij},\ A_{ij} + B_{hk},\ \frac{1}{A_i},\ \frac{A_{ij}}{B_{hk}}.$$

3.4   Prove that if $A_j$ is a type $(0, 1)$ tensor, then

$$\frac{\partial A_j}{\partial x^h} - \frac{\partial A_h}{\partial x^j}$$

is a type $(0, 2)$ tensor.

3.5   (a) If $B_{ij}X^iX^j = 0$ for all $X^i$, show that $B_{ij} = -B_{ji}$.

      (b) If $B_{ij}X^iY^j = 0$ for all $X^i$, $Y^j$, show that $B_{ij} = 0$.

3.6  Show that the equation

$$A_j^h Y_h - \phi Y_j = 0$$

for arbitrary $Y_j$ (where $\phi$ is a scalar), implies that $A_j^h = \phi \delta_j^h$.

3.7  Show that if $a_{hijk} X^h X^j Y^i Y^k = 0$ for all $X^i$, $Y^j$, then

$$a_{hijk} + a_{jihk} + a_{hkji} + a_{jkhi} = 0.$$

3.8  If $B_i{}^h{}_j$ are the components of an arbitrary type $(1, 2)$ tensor and $A^{ijk}$ are such that $A^{ijk} B_i{}^h{}_j$ are the components of a type $(2, 0)$ tensor, show that $A^{ijk}$ are the components of a type $(3, 0)$ tensor.

3.9  If $B^{jk}$ are the components of a skew-symmetric type $(2, 0)$ tensor and $A_{jk}$ are such that $A_{jk} B^{jk}$ is a scalar, show that $(A_{jk} - A_{kj})$ are the components of a type $(0, 2)$ tensor.

3.10  If $bA_{hj} + cA_{jh} = 0$, show that either
(a) $b = -c$ and $A_{hj} = A_{jh}$, or
(b) $b = c$ and $A_{hj} = -A_{jh}$.

3.11  If $A_{hj}$ is skew-symmetric, while $B^{hj}$ is symmetric, show that
(a) $(\delta_l^h \delta_k^j + \delta_k^h \delta_l^j) A_{hj} = 0$,
(b) $A_{hj} B^{hj} = 0$.

3.12  If $B^{jik} = -B^{kji}$ show that $B^{ijk} = 0$. Show that every type $(3, 0)$ tensor which is symmetric on the first pair of indices and skew-symmetric on the last pair of indices vanishes identically.

3.13  Given a tensor $S_{hjk}$, skew-symmetric in $h, j$, find a tensor $T_{hjk}$, skew-symmetric in $j, k$, satisfying the relations

$$-T_{hjk} + T_{jhk} = S_{hjk}.$$

3.14  How many independent components has
(a) $a_{ij}$ if $a_{ij} = a_{ji}$,
(b) $b_{ij}$ if $b_{ij} = -b_{ji}$?

3.15  If $A_{hj}$ is a skew-symmetric tensor of type $(0, 2)$, show that the quantities $B_{rst}$ defined by

$$B_{rst} = \frac{\partial A_{st}}{\partial x^r} + \frac{\partial A_{tr}}{\partial x^s} + \frac{\partial A_{rs}}{\partial x^t}$$

(a) are the components of a tensor; and
(b) are skew-symmetric in all pairs of indices.
(c) Show that $B_{rst}$ has $n(n - 1)(n - 2)/3!$ independent components. What happens if $n = 2$?

3.16  In $X_n$ a tensor with components $H^{ij}{}_{hk}$ has the following properties

$$H^{ij}{}_{hk} = -H^{ji}{}_{hk} = -H^{ij}{}_{kh},$$

$$H^{ij}{}_{hj} = 0.$$

Prove that $H^{ij}{}_{hk}$ has $n^2(n + 1)(n - 3)/4$ independent components. Hence show that if $n = 3$ then

$$H^{ij}{}_{hk} = 0.$$

3.17  If $A^{ijk} = A^{ijk}(x^h)$ is a totally skew-symmetric type $(3, 0)$ tensor field and $a_{ij} = a_{ij}(x^h)$ is a symmetric type $(0, 2)$ tensor field for which $a = \det(a_{ij}) > 0$, show that

$$\frac{1}{\sqrt{a}} \frac{\partial}{\partial x^k} (\sqrt{a} A^{ijk})$$

are the components of a skew-symmetric type $(2, 0)$ tensor.

3.18  Using the transformation law satisfied by the connection coefficients verify by direct calculation that the covariant differential of a type $(1, 2)$ tensor field $T_l{}^j{}_p$, namely,

$$DT_l{}^j{}_p = dT_l{}^j{}_p + \Gamma_m{}^j{}_k T_l{}^m{}_p dx^k - \Gamma_l{}^m{}_k T_m{}^j{}_p dx^k - \Gamma_p{}^m{}_k T_l{}^j{}_m dx^k,$$

is a type $(1, 2)$ tensor.

3.19  Show that the covariant differential of the Kronecker delta vanishes for any connection.

3.20  If $\Gamma_j{}^i{}_k$ are the components of an affine connection show that $\frac{1}{2}(\Gamma_j{}^i{}_k + \Gamma_k{}^i{}_j)$ are also the components of an affine connection.

3.21  If $\phi$ is a scalar field show that

$$\phi_{|i|j} - \phi_{|j|i} = -S_i{}^m{}_j \phi_{|m}.$$

Hence establish that, if the connection is symmetric, so is $\phi_{|i|j}$.

3.22  By considering $A_{i|j} - A_{j|i}$ show that $(\partial A_i/\partial x^j - \partial A_j/\partial x^i)$ are the components of a type $(0, 2)$ tensor. In a similar way show that if $B_{ij} = -B_{ji}$ then $\partial B_{ij}/\partial x^k + \partial B_{ki}/\partial x^j + \partial B_{jk}/\partial x^i$ are the components of a type $(0, 3)$ tensor. (Compare with Problems 3.4 and 3.15.)

3.23  By multiplying (5.17) by $\bar{a}^{iq}(\partial x^m/\partial \bar{x}^q)(\partial x^k/\partial \bar{x}^i)$ and noting (5.10), establish (5.18).

3.24  Show that $\overset{(a)}{\Gamma}_j{}^i{}_k + A_j{}^i{}_k$ are the components of an affine connection.

3.25  If $b_{ij} = b_{ji}$ and $\det(b_{ij}) \neq 0$, show, in the notation of (5.20), that

$$\overset{(a)}{\Gamma}_j{}^i{}_k - \overset{(b)}{\Gamma}_j{}^i{}_k$$

form the components of a type $(1, 2)$ tensor.

3.26  Prove that

(a) $\overset{(a)}{\Gamma}_{hlk} = a_{jl} \overset{(a)}{\Gamma}_h{}^j{}_k;$

(b) $\dfrac{\partial a_{ij}}{\partial x^k} - \dfrac{\partial a_{jk}}{\partial x^i} = \overset{(a)}{\Gamma}_{jik} - \overset{(a)}{\Gamma}_{jki};$

(c) $\dfrac{\partial a^{hj}}{\partial x^k} + a^{mj} \overset{(a)}{\Gamma}_m{}^h{}_k + a^{hm} \overset{(a)}{\Gamma}_m{}^j{}_k = 0.$

3.27 Prove that

$$\frac{\partial(\ln\sqrt{a})}{\partial x^k} = \overset{(a)}{\Gamma}{}^j_{jk},$$

where $a = |\det(a_{hj})|$.

3.28 If $\Gamma^l_{ij}$ are the components of a symmetric affine connection and $a_{ij} = a_{ji}$ is such that $\det(a_{ij}) \neq 0$ and

(a) $a_{ij|k} = 0$

show that

$$\Gamma^l_{ij} = \overset{(a)}{\Gamma}{}^l_{ij};$$

(b) $a_{ij|k} = \lambda_k a_{ij}$

show that

$$\Gamma^l_{ij} = \overset{(a)}{\Gamma}{}^l_{ij} - \tfrac{1}{2}(\lambda_i \delta^l_j + \lambda_j \delta^l_i - a^{lh}\lambda_h a_{ij}).$$

3.29 Show that if

$$\Gamma^l_{ij} = \overset{(a)}{\Gamma}{}^l_{ij}$$

then

$$d(a_{ij}A^i A^j) = 2a_{ij}A^i DA^j.$$

3.30 If $a_{ij}$ is symmetric, $\det(a_{ij}) \neq 0$, and $b_{ij}a_{kh} - b_{ih}a_{jk} + b_{jk}a_{ih} - b_{kh}a_{ij} = 0$ show that $a_{ij} = \lambda b_{ij}$.

3.31 If $a_{ij}$ is symmetric and $\det(a_{ij}) \neq 0$, and

$$\lambda a_{ij}A_{hk} + \mu a_{ih}A_{jk} + \nu a_{ik}A_{hj} = 0$$

show that
　　　(a) if $n > 1$ and $A_{ij} = A_{ji}$,
or　(b) if $n > 2$ and $A_{ij} = -A_{ji}$,
then either $A_{ij} = 0$ or $\lambda = \mu = \nu = 0$.

*3.32 If $a_{ij} = a_{ji}$ and $\det(a_{ij}) \neq 0$, show that $a_{ij}$ are constants under transformations from one coordinate system to another if and only if the transformation is linear. (This result is used in relativity to establish the linearity of the Lorentz transformation of Problem 2.19; see Rindler [1]).

3.33 If $b^{ij}_{|k} = 0$ show that

$$K_m{}^j_{hk}b^{ml} + K_m{}^l_{hk}b^{jm} = 0.$$

3.34 Relative to a rectangular coordinate system in $E_3$ Newton's law of motion is expressed in the form $\mathbf{F} = m\, d\mathbf{v}/dt$, where $\mathbf{F}$ is the force, $m$ the (constant) mass of a particle, and $\mathbf{v}$ the velocity. Express this law in curvilinear coordinates using a covariant differential defined by means of a suitable connection.

3.35 In $X_2$, if $(a_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 x^1 \end{pmatrix}$ compute $\overset{(a)}{\Gamma}{}^i_{jk}$.

Relative to this connection a vector field $X^j$ undergoes parallel displacement along the curve $x^1 = a$ from $x^2 = 0$ to $x^2 = 2\pi$. If initially $X^j = (1, 0)$ calculate $X^j$ finally.

3.36 If $T^{ij}$ is skew-symmetric show that

$$T^{ij}_{\ \ |i|j} = K_{mh}T^{hm} - \tfrac{1}{2}S_h{}^m_k T^{hk}_{\ \ |m}.$$

3.37 If the connection is symmetric, prove that (Veblen [1])

$$K^i_{j\,kl|m} + K^i_{l\,jm|k} + K_m{}^i_{lk|j} + K_k{}^i_{mj|l} = 0.$$

3.38 If $T^j_i$ is a type $(1, 1)$ tensor field show that

$$H_{ij}{}^k = T^r_i \,\partial T^k_j/\partial x^r - T^r_j \,\partial T^k_i/\partial x^r + T^k_r(\partial T^r_i/\partial x^j - \partial T^r_j/\partial x^i)$$

is a type $(1, 2)$ tensor (the Nijenhuis tensor, Nijenhuis [1]).

3.39 Let $a_{ijk}$ be the components of a tensor of type $(0, 3)$. Define

$$S_{ijk} = \frac{1}{3!}(a_{ijk} + a_{kij} + a_{jki} + a_{jik} + a_{kji} + a_{ikj}),$$

$$b_{ijk} = \frac{1}{3!}(a_{ijk} + a_{kij} + a_{jki} - a_{jik} - a_{kji} - a_{ikj}),$$

$$c_{ijk} = \frac{1}{3}(a_{ijk} + a_{jik} - a_{kij} - a_{kji}),$$

$$d_{ijk} = \frac{1}{3}(a_{ijk} + a_{kji} - a_{jik} - a_{jki}).$$

Show that
(a) $S_{ijk}$ is completely symmetric;
(b) $b_{ijk}$ is completely skew-symmetric;
(c) $c_{ijk} = c_{jik}$, $c_{ijk} + c_{kij} + c_{jki} = 0$;
(d) $d_{ijk} = d_{kji}$, $d_{ijk} + d_{kij} + d_{jki} = 0$.
If $P_1 a_{ijk} = S_{ijk}$, $P_2 a_{ijk} = b_{ijk}$, $P_3 a_{ijk} = c_{ijk}$, $P_4 a_{ijk} = d_{ijk}$ show that

$$P_A(P_B a_{ijk}) = (P_A a_{ijk})\delta_{AB} \text{ for } A, B = 1, \ldots, 4 \text{ (no summation over } A).$$

Hence show that $a_{ijk}$ has a unique decomposition of the form

$$a_{ijk} = S_{ijk} + b_{ijk} + c_{ijk} + d_{ijk}.$$

How many independent components has $S_{ijk}, b_{ijk}, c_{ijk},$ and $d_{ijk}$?

3.40 Under the conditions of the previous problem, define

$$h_{ijk} = \tfrac{1}{3}(a_{ijk} - a_{jik} + a_{kji} - a_{kij})$$

$$f_{ijk} = \tfrac{1}{3}(a_{ijk} - a_{kji} + a_{jik} - a_{jki}).$$

Show that (a) $h_{ijk} = -h_{jik}$, $h_{ijk} + h_{kij} + h_{jki} = 0$,

(b) $f_{ijk} = -f_{kji}$, $f_{ijk} + f_{kij} + f_{jki} = 0$.

If $Q_1 a_{ijk} = S_{ijk}$, $Q_2 a_{ijk} = b_{ijk}$, $Q_3 a_{ijk} = h_{ijk}$, $Q_4 a_{ijk} = f_{ijk}$ show that $Q_A(Q_B a_{ijk})$ $= (Q_A a_{ijk})\delta_{AB}$, $(A, B = 1, \ldots, 4)$. Hence show that $a_{ijk}$ has a unique decomposition of the form

$$a_{ijk} = S_{ijk} + b_{ijk} + h_{ijk} + f_{ijk}.$$

How many independent components has $h_{ijk}$ and $f_{ijk}$?

[Problems 3.39 and 3.40 may be regarded as two distinct generalizations of the decomposition (2.18) to tensors of type (0, 3).]

3.41  By differentiating (2.23) with respect to $X^a$, obtain (2.25).

3.42  Obtain (3.21) from (3.16) by interchanging $x^i$ with $\bar{x}^i$.

3.43  Obtain (8.6) from (6.10) by setting $Y_j = \phi_{|j}$ and using Problem 3.21.

3.44  Show that the coefficients of the affine connection obtained by first transforming from $x^i$ to $\bar{x}^i$ by $\bar{x}^i = \bar{x}^i(x^j)$ and then from $\bar{x}^i$ to $\bar{\bar{x}}^i$ by $\bar{\bar{x}}^i = \bar{\bar{x}}^i(\bar{x}^j)$ are identical with those coefficients obtained by transforming directly from $x^i$ to $\bar{\bar{x}}^i$ by means of $\bar{\bar{x}}^i = \bar{\bar{x}}^i(\bar{x}^j(x^h))$.

3.45  In $X_4$, Newton's equations of motion of a particle in a gravitational field characterized by a scalar field $\phi = \phi(x^i)$ are

$$\frac{d^2 x^4}{dt^2} = 0, \qquad \frac{d^2 x^\alpha}{dt^2} = -\frac{\partial \phi}{\partial x^\alpha} \qquad (\alpha = 1, 2, 3).$$

Show that these are autoparallel curves of $X_4$ if the nonzero components of the connection are $\Gamma_4{}^\alpha{}_4 = a^2 \, \partial\phi/\partial x^\alpha$, where $a$ is a constant. Hence show that the only nonzero components of the curvature tensor are $K_4{}^\alpha{}_{4\beta} = a^2 \, \partial^2\phi/\partial x^\alpha \, \partial x^\beta$ and that $K_{ij} = 0$ is equivalent to Laplace's equation. By considering $a^{ij}{}_{|k}$ show that $\Gamma_j{}^i{}_k$ are Christoffel symbols of the second kind with respect to the symmetric tensor field $a_{ij}$ if and only if $X_4$ is flat. (The geometry of dynamics has been treated from a tensorial point of view with the aid of a symmetric type (0, 2) tensor field in great detail in distinct ways by Synge [1].)

# 4

# ADDITIONAL TOPICS FROM THE TENSOR CALCULUS

In this chapter we discuss several somewhat disjoint topics which, despite their importance, do not fit naturally into the development of the previous chapter. An important extension of the tensor concept is represented by the so-called relative tensors; although this generalization is a fairly simple one which merely entails some minor modifications of the general theory, it is indispensable from the point of view of physical applications. This is due to the fact that the integral of a scalar is not itself a scalar; indeed, in order that an $n$-fold integral over a region of $X_n$ be invariant under arbitrary coordinate transformations it is necessary and sufficient that its integrand be a relative scalar (also called a scalar density).

Some considerable space is also devoted to the numerical relative tensors (of which the Kronecker delta is a special case); intrinsically these are not particularly interesting, but, as will be seen in subsequent chapters, they are exceedingly useful and powerful manipulative tools. A brief description is also given of normal coordinates, which, in a sense, may be regarded as generalizations of spherical polar coordinates of Euclidean geometry. Sometimes the application of normal coordinates gives rise to useful simplifications of calculations involving tensors; however, their use is fraught with danger. A very brief discussion is also given of the so-called Lie derivatives which appear frequently in the modern literature. In the present context such derivatives are defined with the aid of an arbitrary vector field, without reference to an underlying connection. An alternative approach to the Lie derivative is sketched in the Appendix.

## 4.1 RELATIVE TENSORS

In this section we introduce a generalization of the tensor concept which occurs very frequently in applications, particularly when integration processes are involved (Synge and Schild [1]). Although this extension is important, it is not an unduly drastic one; in fact, it will be seen that it is a simple matter to apply the techniques and results of the previous chapter to the generalization to be considered below.

We shall begin with an example which is suggested by a type $(0, 2)$ tensor field $a_{hk}(x^p)$ defined on our differentiable manifold $X_n$. Under a class $C^2$ coordinate transformation

$$\bar{x}^j = \bar{x}^j(x^h), \tag{1.1}$$

the transformed components of $a_{hk}$ are given by

$$\bar{a}_{jl} = \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial x^k}{\partial \bar{x}^l} a_{hk}. \tag{1.2}$$

According to the usual product rule for determinants one therefore finds that

$$\det(\bar{a}_{jl}) = \det\left(\frac{\partial x^h}{\partial \bar{x}^j}\right)\det\left(\frac{\partial x^k}{\partial \bar{x}^l}\right)\det(a_{hk}). \tag{1.3}$$

We shall write, for the sake of brevity,

$$a(x^p) = \det[a_{hk}(x^p)], \tag{1.4}$$

while in accordance with (3.1.6) the Jacobian of the inverse of (1.1) is denoted by $J$:

$$J = \det\left(\frac{\partial x^h}{\partial \bar{x}^j}\right) = \frac{\partial(x^1, \ldots, x^n)}{\partial(\bar{x}^1, \ldots, \bar{x}^n)}, \tag{1.5}$$

of which it is assumed that it is positive. Thus (1.3) becomes

$$\bar{a}(\bar{x}^m) = J^2 a(x^p). \tag{1.6}$$

This result indicates very clearly that the determinant of a type $(0, 2)$ tensor is *not* a scalar, nor is it a tensor of any type $(r, s)$, with $r, s > 0$. In fact, it is evident that (1.6) exemplifies a transformation law which differs from those encountered thus far. Furthermore, if we assume for the sake of the present argument that $a(x^p)$ is nonnegative, it follows from (1.6) that this is true also of $\bar{a}(\bar{x}^m)$, and accordingly it may be inferred from (1.6) that

$$[\bar{a}(\bar{x}^m)]^{1/2} = J[a(x^p)]^{1/2}. \tag{1.7}$$

This is again a new transformation law; further transformation laws may be obtained by taking different exponents of (1.6).

Clearly the examples (1.6) and (1.7) are nontrivial, and accordingly we are led to introduce the following.

### DEFINITION

*A function $\psi(x^h)$ of the coordinates of the manifold $X_n$ is said to represent a relative scalar field of weight w, if, under the coordinate transformation (1.1), the transform of this function is given by*

$$\bar{\psi}(\bar{x}^j) = J^w \psi(x^h). \tag{1.8}$$

*Remark.* A relative scalar of unit weight is often called a *scalar density*.

Accordingly the determinant (1.4) is a relative scalar of weight 2, while its square root is a relative scalar of weight 1, as indicated respectively by (1.6) and (1.7). Furthermore, it is obvious that a scalar (or invariant) as defined in Section 3.2 is simply a relative scalar of weight 0.

The importance of the concept of scalar density is illustrated by the following consideration. Let us suppose that we are given a closed, simply connected region $G$ in $X_n$, on which a continuous function $f(x^h)$ of the coordinates is defined, so that we may construct the $n$-fold integral

$$I = \int_G f(x^h)\, dx^1 \cdots dx^n. \tag{1.9}$$

[For instance, an integral of this type would define the mass of a material body which occupies the region $G$ if $f(x^h)$ denotes the density of matter.] According to the usual rule for a change of the independent variables in the integrand of a multiple integral (Apostol [1]) it is found that the value of the integral corresponding to (1.9) in the $\bar{x}$-coordinate system is given by

$$\bar{I} = \int_G \bar{f}(\bar{x}^j)\, d\bar{x}^1 \cdots d\bar{x}^n = \int_G \bar{f}(\bar{x}^j)\left|\frac{\partial(\bar{x}^1, \ldots, \bar{x}^n)}{\partial(x^1, \ldots, x^n)}\right|\, dx^1 \cdots dx^n,$$

where $\bar{f}(\bar{x}^j)$ is the transform of $f(x^h)$. By means of (1.5) this can be written as

$$\bar{I} = \int_G \bar{f}(\bar{x}^j) J^{-1}\, dx^1 \cdots dx^n. \tag{1.10}$$

If $f(x^h)$ is a scalar, that is, if $\bar{f}(\bar{x}^j) = f(x^h)$, it would follow immediately from (1.9) and (1.10) that $\bar{I} \neq I$ (unless $J = 1$, which is not assumed here). Thus the integral of a scalar field is not, in general, a scalar. However, it is frequently of utmost importance to construct invariant integrals: in fact, the theory of most physical fields, such as the electromagnetic field, or the gravitational field of general relativity, depends crucially on the construction of invariant integrals. (Also, the simple example referred to above is indicative of

this requirement, for surely the mass of a material body is independent of the choice of special coordinates.)

On the other hand, if the integrand $f(x^h)$ of (1.9) happens to be a scalar density, it follows immediately from (1.8) that

$$\bar{f}(\bar{x}^j)J^{-1} = f(x^h),$$

and if this is substituted in (1.10), one obtains

$$\int_G \bar{f}(\bar{x}^j) \, d\bar{x}^1 \cdots d\bar{x}^n = \int_G f(x^h) \, dx^1 \cdots dx^n, \qquad (1.11)$$

so that the required invariance of the integral (1.9) is indeed attained. We therefore conclude that *a necessary and sufficient condition in order that the integral* (1.9) *be invariant under the coordinate transformation* (1.1) *is that the integrand $f(x^h)$ be a scalar density.*

Thus far we have merely generalized the concept of scalar, and the question arises as to whether a corresponding extension in respect of tensors of arbitrary type $(r, s)$ is feasible. Again, the example of our type $(0, 2)$ tensor field provides us with an immediate motivation. Let us denote the cofactor of the element $a_{hk}$ in $a = \det(a_{hk})$ by $A^{hk}$, so that, according to (1.3.10), the elements of the inverse matrix $(a^{hk})$ are given by

$$aa^{hk} = A^{hk}. \qquad (1.12)$$

But we know that the $a^{hk}$ constitute the components of a type $(2, 0)$ tensor. Thus if we apply the transformation law of the latter, together with (1.6), to (1.12) we see that

$$\bar{A}^{lj} = \bar{a}\bar{a}^{lj} = J^2 a \frac{\partial \bar{x}^l}{\partial x^h} \frac{\partial \bar{x}^j}{\partial x^k} a^{hk} = J^2 \frac{\partial \bar{x}^l}{\partial x^h} \frac{\partial \bar{x}^j}{\partial x^k} A^{hk}. \qquad (1.13)$$

It follows that the cofactors $A^{hk}$ in $\det(a_{hk})$ do not satisfy the transformation law of a type $(2, 0)$ tensor; again the factor $J^2$ on the right-hand side of (1.13) is responsible for this. This situation is analogous to (1.6), and accordingly we shall refer to quantities which transform in the manner indicated by (1.13) as relative type $(2, 0)$ tensors of weight 2.

The above considerations clearly indicate the need for the following general definition.

### DEFINITION

*A set of $n^{r+s}$ functions $\Lambda^{h_1 \cdots h_r}{}_{k_1 \cdots k_s}(x^q)$ is said to constitute the components of a relative tensor field of type $(r, s)$ and weight $w$ on the manifold $X_n$, if, under the coordinate transformation* (1.1), *these functions transform according to the*

*relation*

$$\bar{\Lambda}^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}(\bar{x}^p) = J^w \frac{\partial \bar{x}^{j_1}}{\partial x^{h_1}} \cdots \frac{\partial \bar{x}^{j_r}}{\partial x^{h_r}} \cdot \frac{\partial x^{k_1}}{\partial \bar{x}^{l_1}} \cdots \frac{\partial x^{k_s}}{\partial \bar{x}^{l_s}} \Lambda^{h_1 \cdots h_r}{}_{k_1 \cdots k_s}(x^q), \quad (1.14)$$

*in which the Jacobian $J$ is defined by* (1.5).

*Remark 1.* For the purposes of this definition the weight $w$ is restricted to the values $0, \pm 1, \pm 2, \ldots$.

*Remark 2.* Relative tensors of unit weight are generally called tensor *densities.* In order to distinguish the usual concept of tensor [$w = 0$ in (1.14)], relative tensors of zero weight are sometimes referred to as *absolute* tensors.

*Remark 3.* The algebraic operations described in Section 3.2 in respect of tensors are easily extended to the case of relative tensors:

1.   Two relative tensors of identical type and weight may be added to yield a relative tensor of the same type and weight.
2.   The products of the components of a type $(r_1, s_1)$ relative tensor of weight $w_1$ with those of a type $(r_2, s_2)$ relative tensor of weight $w_2$ constitute the components of a type $(r_1 + r_2, s_1 + s_2)$ relative tensor of weight $w_1 + w_2$; in particular, when a type $(r, s)$ relative tensor of weight $w_1$ is multiplied by a relative scalar of weight $w_2$, a type $(r, s)$ relative tensor of weight $w_1 + w_2$ is obtained.
3.   The process of contraction applied to a type $(r, s)$ relative tensor of weight $w$ yields a type $(r - 1, s - 1)$ relative tensor of the same weight (provided that $r \geq 1, s \geq 1$).
4.   Symmetry and skew-symmetry properties of relative tensors of arbitrary weight are independent of the choice of the coordinate system.

These statements are immediate consequences of the properties of the corresponding operations described in Section 3.2. Furthermore, from the linearity of (1.14) in the components of the relative tensor field it follows that if a relative tensor vanishes in a given coordinate system, this will be the case also in all other coordinate systems. *Accordingly any relation which is expressed entirely in terms of relative tensors is invariant under coordinate transformations.*

Again, the ordinary partial derivatives of a relative tensor field do not constitute the components of a relative tensor, and accordingly it is necessary to define once more a process of covariant differentiation. It will be seen that this requires a nontrivial modification of the method introduced in the preceding chapter. This situation is best illustrated by the consideration of the gradient of a relative scalar field of weight $w$. To this end we differentiate (1.8)

with respect to $\bar{x}^j$, which gives

$$\frac{\partial \overline{\psi}}{\partial \bar{x}^j} = J^w \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \psi}{\partial x^h} + w J^{w-1} \frac{\partial J}{\partial \bar{x}^j} \psi. \tag{1.15}$$

In view of the presence of the second term on the right-hand side it is obvious that the gradient of a scalar density of weight $w$ is not a relative covariant vector (unless $w = 0$, which is not assumed here). In order to construct an appropriate relative tensor we have to evaluate explicitly the partial derivatives of the Jacobian $J$.

Let us denote the cofactor of the element $\partial x^h / \partial \bar{x}^l$ of $J$ by $C_h^l$, so that

$$\frac{\partial x^h}{\partial \bar{x}^l} C_k^l = J \delta_k^h. \tag{1.16}$$

But from (3.1.3) we have

$$\frac{\partial x^h}{\partial \bar{x}^l} \cdot \frac{\partial \bar{x}^l}{\partial x^k} = \delta_k^h, \tag{1.17}$$

so that

$$C_k^l = J \frac{\partial \bar{x}^l}{\partial x^k}. \tag{1.18}$$

Now, according to the usual rule for the differentiation of a determinant as given in Section 1.3, we have

$$\frac{\partial J}{\partial \bar{x}^j} = \left[ \frac{\partial}{\partial \bar{x}^j} \left( \frac{\partial x^h}{\partial \bar{x}^l} \right) \right] \cdot C_h^l,$$

so that, by (1.18),

$$\frac{\partial J}{\partial \bar{x}^j} = J \frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l} \frac{\partial \bar{x}^l}{\partial x^h}. \tag{1.19}$$

This is substituted in (1.15) to yield

$$\frac{\partial \overline{\psi}}{\partial \bar{x}^j} = J^w \left[ \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \psi}{\partial x^h} + w \frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l} \frac{\partial \bar{x}^l}{\partial x^h} \psi \right], \tag{1.20}$$

which represents the explicit transformation law of the gradient of $\psi$.

In order to construct a suitable tensor density, we now invoke the transformation law (3.3.21) of the connection coefficients, which is

$$\frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l} \frac{\partial \bar{x}^m}{\partial x^h} = \overline{\Gamma}_{j\,l}^{\,m} - \frac{\partial \bar{x}^m}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial x^k}{\partial \bar{x}^l} \Gamma_{h\,k}^{\,p}.$$

In order to be able to substitute this in (1.20), we must contract over the indices $m$ and $l$; with the aid of (1.17) we thus obtain

$$\frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l} \frac{\partial \bar{x}^l}{\partial x^h} = \overline{\Gamma}_{j\,l}^{\,l} - \delta_p^k \frac{\partial x^h}{\partial \bar{x}^j} \Gamma_{h\,k}^{\,p} = \overline{\Gamma}_{j\,l}^{\,l} - \frac{\partial x^h}{\partial \bar{x}^j} \Gamma_{h\,k}^{\,k}. \tag{1.21}$$

This, incidentally, is the transformation law of the contracted connection coefficients. If this is substituted in (1.20), it is found that

$$\frac{\partial \overline{\psi}}{\partial \bar{x}^j} = J^w \frac{\partial x^h}{\partial \bar{x}^j} \frac{\partial \psi}{\partial x^h} + w J^w \overline{\Gamma}_{j\,l}^{\,l} \psi - w J^w \frac{\partial x^h}{\partial \bar{x}^j} \Gamma_{h\,k}^{\,k} \psi,$$

or, if we apply (1.8) to the second term on the right-hand side,

$$\frac{\partial \overline{\psi}}{\partial \bar{x}^j} - w \overline{\Gamma}_{j\,l}^{\,l} \overline{\psi} = J^w \frac{\partial x^h}{\partial \bar{x}^j} \left[ \frac{\partial \psi}{\partial x^h} - w \Gamma_{h\,k}^{\,k} \psi \right], \tag{1.22}$$

from which it is immediately evident that the quantities defined by

$$\psi_{|h} = \frac{\partial \psi}{\partial x^h} - w \Gamma_{h\,k}^{\,k} \psi, \tag{1.23}$$

constitute the components of a type $(0, 1)$ relative tensor of weight $w$. It is therefore obvious that, when the covariant derivative of a relative scalar is constructed, an additional term involving the contracted connection coefficients has to be introduced.

By means of (1.21) the argument given above is easily extended to type $(r, s)$ relative tensor fields of weight $w$, and it is found that *the covariant derivative of such a field must be defined as follows:*

$$\Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_s | k} = \frac{\partial \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}}{\partial x^k} + \sum_{\alpha=1}^{r} \Gamma_{m\,k}^{\,j_\alpha} \Lambda^{j_1 \cdots j_{\alpha-1} m j_{\alpha+1} \cdots j_r}{}_{l_1 \cdots l_s}$$
$$- \sum_{\beta=1}^{s} \Gamma_{l_\beta\,k}^{\,m} \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_{\beta-1} m l_{\beta+1} \cdots l_s} - w \Gamma_{k\,h}^{\,h} \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}, \tag{1.24}$$

which obviously reduces to the definition (3.5.7) of the covariant derivative of a type $(r, s)$ tensor field whenever $w = 0$.

For example, the covariant derivative of a relative contravariant vector field of weight $w$ is

$$A^j_{|k} = \frac{\partial A^j}{\partial x^k} + \Gamma_{h\,k}^{\,j} A^h - w \Gamma_{k\,h}^{\,h} A^j. \tag{1.25}$$

In particular, if we contract over the indices $j$ and $k$ in this relation, we obtain

$$A^j_{|j} = \frac{\partial A^j}{\partial x^j} + \Gamma_{h\,j}^{\,j} A^h - w \Gamma_{j\,h}^{\,h} A^j = \frac{\partial A^j}{\partial x^j} + (1 - w) \Gamma_{h\,j}^{\,j} A^h. \tag{1.26}$$

Thus for a relative contravariant vector field of unit weight the contracted covariant derivative is simply the sum of the corresponding ordinary partial derivatives:

$$A^j_{|j} = \frac{\partial A^j}{\partial x^j} \qquad (w = 1). \tag{1.27}$$

In analogy with the corresponding concept of elementary vector analysis we regard $A^j_{|j}$ (for *any* value of $w$) as the *divergence* of the relative contravariant vector field $A^j$.

Since the contracted connection coefficients obviously play an important role whenever tensor densities are considered, let us evaluate these coefficients explicitly for the case when the connection is derived directly from a non-singular class $C^1$ symmetric type (0, 2) tensor field $a_{hk}(x^l)$ as was done in Section 3.5. From (3.5.15) and (3.5.20) we have, for the Christoffel symbols which define such a connection,

$$\overset{(a)}{\Gamma}_{h}{}^{j}{}_{k} = \frac{1}{2} a^{lj}\left(\frac{\partial a_{kl}}{\partial x^h} + \frac{\partial a_{lh}}{\partial x^k} - \frac{\partial a_{hk}}{\partial x^l}\right). \tag{1.28}$$

Contracting over $j$ and $k$ in this relation, we obtain

$$\overset{(a)}{\Gamma}_{h}{}^{j}{}_{j} = \frac{1}{2} a^{lj}\left(\frac{\partial a_{jl}}{\partial x^h} + \frac{\partial a_{lh}}{\partial x^j} - \frac{\partial a_{hj}}{\partial x^l}\right). \tag{1.29}$$

But as a result of the postulated symmetry of $a_{hj}$ we have

$$a^{lj}\frac{\partial a_{lh}}{\partial x^j} = a^{jl}\frac{\partial a_{jh}}{\partial x^l} = a^{lj}\frac{\partial a_{hj}}{\partial x^l},$$

so that (1.29) reduces to

$$\overset{(a)}{\Gamma}_{h}{}^{j}{}_{j} = \frac{1}{2} a^{jl}\frac{\partial a_{jl}}{\partial x^h} = \frac{1}{2} a^{-1} A^{jl}\frac{\partial a_{jl}}{\partial x^h}, \tag{1.30}$$

where, in the last step, we have used (1.12). Accordingly (1.30) becomes

$$\overset{(a)}{\Gamma}_{h}{}^{j}{}_{j} = \frac{1}{2} a^{-1}\frac{\partial a}{\partial x^h} = \frac{1}{2}\frac{\partial}{\partial x^h}(\ln a) = \frac{\partial}{\partial x^h}(\ln \sqrt{a}) = \frac{1}{\sqrt{a}}\frac{\partial \sqrt{a}}{\partial x^h}, \tag{1.31}$$

in which $a$ denotes $|\det(a_{hj})|$.

*This is the required expression for the contracted Christoffel symbols.*

Incidentally, since $\sqrt{a}$ is a scalar density of unit weight in view of (1.7), it follows from (1.23) that

$$(\sqrt{a})_{|h} = \frac{\partial \sqrt{a}}{\partial x^h} - \Gamma_{h}{}^{j}{}_{j}\sqrt{a} \tag{1.32}$$

for an arbitrary connection. However, if we choose the special connection (1.28) it follows immediately from (1.31) that

$$(\sqrt{a})_{|h} = 0, \tag{1.33}$$

which is a natural counterpart of the identities (3.5.26) and (3.5.27).

## 4.2   THE NUMERICAL RELATIVE TENSORS

In this section we are concerned with the so-called numerical relative tensors which are often extremely useful in complicated manipulative processes involving tensors (Veblen [2]). We have already encountered an example of a numerical tensor, namely, the Kronecker delta. This tensor is a special case of one of the most important numerical tensors, the so-called generalized Kronecker delta. The *generalized Kronecker delta*, denoted by $\delta^{j_1 \cdots j_r}_{h_1 \cdots h_r}$, possesses an equal number of sub- and superscripts ($r$) and is defined in terms of an $r \times r$ determinant as follows:

$$\delta^{j_1\cdots j_r}_{h_1\cdots h_r} = \begin{vmatrix} \delta^{j_1}_{h_1} & \delta^{j_1}_{h_2} & \cdots & \delta^{j_1}_{h_r} \\ \delta^{j_2}_{h_1} & \delta^{j_2}_{h_2} & \cdots & \delta^{j_2}_{h_r} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \delta^{j_r}_{h_1} & \delta^{j_r}_{h_2} & \cdots & \delta^{j_r}_{h_r} \end{vmatrix}. \tag{2.1}$$

Clearly the Kronecker delta $\delta^j_h$ is a special case of this definition which corresponds to the value $r = 1$. With $r = 2$ in (2.1) we see that

$$\delta^{jh}_{ik} = \delta^j_i\delta^h_k - \delta^j_k\delta^h_i.$$

In general $\delta^{j_1\cdots j_r}_{h_1\cdots h_r}$ is the sum of $r!$ terms, each of which is the product of $r$ Kronecker deltas. Since, as we have seen, the ordinary Kronecker delta is a type (1, 1) tensor, it follows immediately that *the generalized Kronecker delta (2.1) is a type $(r, r)$ tensor.*

In (2.1) we notice that the indices $j_1, j_2, \ldots, j_r$ correspond to the 1st, 2nd, ..., $r$th rows of the determinant, whereas $h_1, h_2, \ldots, h_r$ correspond to the 1st, 2nd, ..., $r$th columns of the determinant. In view of the fact that the sign of a determinant is changed by interchanging two rows (or two columns) we immediately have the following important properties:

$$\begin{aligned} \delta^{i_1\cdots i_h\cdots i_k\cdots i_r}_{j_1\quad\cdots\quad\ j_r} &= -\delta^{i_1\cdots i_k\cdots i_h\cdots i_r}_{j_1\quad\cdots\quad\ j_r}, \\ \delta^{i_1\quad\cdots\quad\ i_r}_{j_1\cdots j_h\cdots j_k\cdots j_r} &= -\delta^{i_1\quad\cdots\quad\ i_r}_{j_1\cdots j_k\cdots j_h\cdots j_r}, \end{aligned} \tag{2.2}$$

for all distinct $h, k$, $1 \le h, k \le r$; that is, the generalized Kronecker delta $\delta^{i_1\cdots i_r}_{j_1\cdots j_r}$ is skew-symmetric under interchange of any two of the indices $i_1 \cdots i_r$

and is skew-symmetric under interchange of any two of the indices $j_1 \cdots j_r$. Clearly then, if any two of the indices $i_1 \cdots i_r$ (or $j_1 \cdots j_r$) coincide, the corresponding $\delta^{i_1 \cdots i_r}_{j_1 \cdots j_r}$ vanishes identically. If $r > n$, then at least two of the indices $i_1 \cdots i_r$ must coincide, in which case we have the useful result

$$\delta^{i_1 \cdots i_r}_{j_1 \cdots j_r} \equiv 0 \qquad \text{if } r > n. \tag{2.3}$$

To demonstrate the application of (2.3), and to gain some manipulative skill involving (2.1), we briefly consider two examples, which will be used in the sequel.

EXAMPLE 1    CURVATURE TENSOR IN $X_2$

For $n = 2$ we have, from (2.3) and (2.1),

$$0 = \delta^{ijk}_{lst} = \delta^k_t \delta^{ij}_{ls} - \delta^k_s \delta^{ij}_{lt} + \delta^k_l \delta^{ij}_{st}.$$

When the latter is multiplied by $K_h{}^l{}_{ij}$, the curvature tensor of $X_2$ defined by (3.6.8), account being taken of (3.8.1), we find

$$0 = -2\delta^k_t K_h{}^l{}_{sl} + 2\delta^k_s K_h{}^l{}_{tl} + 2K_h{}^k{}_{st}. \tag{2.4}$$

By virtue of (3.8.13), (2.4) can be expressed in the form

$$K_h{}^k{}_{st} = \delta^k_t K_{hs} - \delta^k_s K_{ht} \qquad \text{for } n = 2; \tag{2.5}$$

that is, in $X_2$ the Ricci tensor completely determines the curvature tensor.

EXAMPLE 2    TENSORS OF TYPE (2, 2) IN $X_3$

Consider a tensor of type (2, 2), $H^{ij}{}_{hk}$, which has the following properties:

$$H^{ij}{}_{hk} = -H^{ji}{}_{hk} = -H^{ij}{}_{kh},$$
$$H^{ij}{}_{hj} = 0. \tag{2.6}$$

(Specific examples of tensors with these properties are discussed later.) We now wish to evaluate the tensor $\delta^{klrs}_{ijtu} H^{ij}{}_{kl}$. Expanding $\delta^{klrs}_{ijtu}$ about the first row (the index $k$) we see that, by (2.6),

$$\delta^{klrs}_{ijtu} H^{ij}{}_{kl} = \delta^{lrs}_{iju} H^{ij}{}_{tl} - \delta^{lrs}_{ijt} H^{ij}{}_{ul}.$$

Each of the generalized Kronecker deltas on the right-hand side of the latter are now expanded about the first row (the index $l$) so that, by (2.6),

$$\delta^{klrs}_{ijtu} H^{ij}{}_{kl} = \delta^{rs}_{ij} H^{ij}{}_{tu} - \delta^{rs}_{ij} H^{ij}{}_{ut},$$

which finally yields

$$\delta^{klrs}_{ijtu} H^{ij}{}_{kl} = 4H^{rs}{}_{tu}. \tag{2.7}$$

Since, for $n = 3$, (2.3) implies

$$\delta^{klrs}_{ijtu} = 0$$

we see from (2.7) that if $H^{ij}{}_{kl}$ satisfies (2.6) then

$$H^{ij}{}_{kl} \equiv 0 \qquad \text{for } n = 3. \tag{2.8}$$

In order to obtain other important properties of the generalized Kronecker delta, we now return to (2.1) and, by virtue of (2.3), we restrict ourselves to the case $r \le n$. Let us expand (2.1) in terms of the elements of its last column and the cofactors thereof, the latter being $(r - 1) \times (r - 1)$ determinants, which are again expressed by means of the generalized Kronecker delta. One thus obtains

$$\delta^{j_1 \cdots j_r}_{h_1 \cdots h_r} = \delta^{j_r}_{h_r} \delta^{j_1 \cdots j_{r-1}}_{h_1 \cdots h_{r-1}} - \delta^{j_{r-1}}_{h_r} \delta^{j_1 \cdots j_{r-2} \, j_r}_{h_1 \cdots \cdots h_{r-1}} + \delta^{j_{r-2}}_{h_r} \delta^{j_1 \cdots j_{r-3} j_{r-1} j_r}_{h_1 \cdots \cdots h_{r-1}}$$
$$+ \cdots + (-1)^{r+1} \delta^{j_1}_{h_r} \delta^{j_2 \cdots j_r}_{h_1 \cdots h_{r-1}}. \tag{2.9}$$

We now contract over the indices $j_r$, $h_r$, obtaining

$$\delta^{j_1 \cdots j_{r-1} \, j_r}_{h_1 \cdots h_{r-1} \, j_r} = n\delta^{j_1 \cdots j_{r-1}}_{h_1 \cdots h_{r-1}} - \delta^{j_1 \cdots j_{r-2} \, j_{r-1}}_{h_1 \cdots \cdots h_{r-1}} + \delta^{j_1 \cdots j_{r-3} j_{r-1} j_{r-2}}_{h_1 \cdots \cdots h_{r-1}}$$
$$+ \cdots + (-1)^{r+1} \delta^{j_2 \cdots j_{r-1} j_1}_{h_1 \cdots \cdots h_{r-1}}$$

However, by virtue of (2.2), the latter reduces to

$$\delta^{j_1 \cdots j_{r-1} j_r}_{h_1 \cdots h_{r-1} j_r} = (n - r + 1)\delta^{j_1 \cdots j_{r-1}}_{h_1 \cdots h_{r-1}}. \tag{2.10}$$

In this formula we now contract over $j_{r-1}$, $h_{r-1}$, applying the formula to itself for the case when $r$ is reduced to $r - 1$. This gives

$$\delta^{j_1 \cdots j_{r-2} j_{r-1} j_r}_{h_1 \cdots h_{r-2} j_{r-1} j_r} = (n - r + 2)(n - r + 1)\delta^{j_1 \cdots j_{r-2}}_{h_1 \cdots h_{r-2}}.$$

If this process is repeated, we obtain the following general identity:

$$\delta^{j_1 \cdots j_s j_{s+1} \cdots j_r}_{h_1 \cdots h_s j_{s+1} \cdots j_r} = \frac{(n - s)!}{(n - r)!} \delta^{j_1 \cdots j_s}_{h_1 \cdots h_s}, \tag{2.11}$$

which is frequently used. An equivalent form of (2.11) is the following:

$$\delta^{j_1 \cdots j_t j_{t+1} \cdots j_r}_{j_1 \cdots j_t h_{t+1} \cdots h_r} = \frac{(n - r + t)!}{(n - r)!} \delta^{j_{t+1} \cdots j_r}_{h_{t+1} \cdots h_r}. \tag{2.12}$$

In particular, when $t = r$, we have

$$\delta^{j_1 \cdots j_r}_{j_1 \cdots j_r} = \frac{n!}{(n - r)!}. \tag{2.13}$$

As a further example, let us consider the sum $\delta^{jhk}_{lst} A_{jhk}$, where $A_{jhk}$ is a type (0, 3) tensor which is supposed to be skew-symmetric in all its indices.

By (2.9) (with $r = 3$) we have

$$\delta^{jhk}_{lst} A_{jhk} = \delta^{jh}_{ls} A_{jht} - \delta^{jk}_{ls} A_{jtk} + \delta^{hk}_{ls} A_{thk}$$
$$= 3! A_{lst}.$$

In fact, for any integer $r \le n$, it is easily seen that

$$\delta^{j_1 \cdots j_r}_{h_1 \cdots h_r} A_{j_1 j_2 \cdots j_r} = r! A_{h_1 h_2 \cdots h_r}, \tag{2.14}$$

provided that $A_{j_1 \cdots j_r}$ is skew-symmetric in all its indices. In a similar way we also have

$$\delta^{j_1 \cdots j_r}_{h_1 \cdots h_r} B^{h_1 \cdots h_r} = r! B^{j_1 \cdots j_r}, \tag{2.15}$$

provided that $B^{h_1 \cdots h_r}$ is skew-symmetric in all its indices.

We shall now derive another important result associated with the generalized Kronecker delta. Let $T_{i_1 \cdots i_r}$ be a type $(0, r)$ tensor field so that

$$\overline{T}_{i_1 \cdots i_r}(\overline{x}^l) = \frac{\partial x^{h_1}}{\partial \overline{x}^{i_1}} \cdots \frac{\partial x^{h_r}}{\partial \overline{x}^{i_r}} T_{h_1 \cdots h_r}(x^k).$$

We differentiate this equation with respect to $\overline{x}^{i_{r+1}}$ to find

$$\frac{\partial \overline{T}_{i_1 \cdots i_r}}{\partial \overline{x}^{i_{r+1}}} = \sum_{\mu=1}^{r} \frac{\partial^2 x^{h_\mu}}{\partial \overline{x}^{i_{r+1}} \partial \overline{x}^{i_\mu}} \frac{\partial x^{h_1}}{\partial \overline{x}^{i_1}} \cdots \frac{\partial x^{h_{\mu-1}}}{\partial \overline{x}^{i_{\mu-1}}} \frac{\partial x^{h_{\mu+1}}}{\partial \overline{x}^{i_{\mu+1}}} \cdots \frac{\partial x^{h_r}}{\partial \overline{x}^{i_r}} T_{h_1 \cdots h_r}$$
$$+ \frac{\partial x^{h_1}}{\partial \overline{x}^{i_1}} \cdots \frac{\partial x^{h_r}}{\partial \overline{x}^{i_r}} \frac{\partial x^{h_{r+1}}}{\partial \overline{x}^{i_{r+1}}} \frac{\partial T_{h_1 \cdots h_r}}{\partial x^{h_{r+1}}}, \tag{2.16}$$

which clearly indicates that $\partial T_{h_1 \cdots h_r} / \partial x^{h_{r+1}}$ are not tensors, in accord with Section 3.5. However, the generalized Kronecker delta enables us to associate with the latter a quantity which is tensorial. This is achieved in the following way. If we multiply (2.16) by $\overline{\delta}^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}}$ and note that $\partial^2 x^{h_\mu} / \partial \overline{x}^{i_{r+1}} \partial \overline{x}^{i_\mu}$ is symmetric in $i_{r+1} i_\mu$ for $\mu = 1, \ldots, r$ whereas $\overline{\delta}^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}}$ is skew-symmetric in $i_{r+1} i_\mu$ for $\mu = 1, \ldots, r$ we thus find

$$\overline{\delta}^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}} \frac{\partial \overline{T}_{i_1 \cdots i_r}}{\partial \overline{x}^{i_{r+1}}} = \overline{\delta}^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}} \frac{\partial x^{h_1}}{\partial \overline{x}^{i_1}} \cdots \frac{\partial x^{h_{r+1}}}{\partial \overline{x}^{i_{r+1}}} \frac{\partial T_{h_1 \cdots h_r}}{\partial x^{h_{r+1}}}. \tag{2.17}$$

Since $\delta^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}}$ is a type $(r + 1, r + 1)$ tensor we also have

$$\overline{\delta}^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}} \frac{\partial x^{h_1}}{\partial \overline{x}^{i_1}} \cdots \frac{\partial x^{h_{r+1}}}{\partial \overline{x}^{i_{r+1}}} = \delta^{h_1 \cdots h_{r+1}}_{k_1 \cdots k_{r+1}} \frac{\partial x^{k_1}}{\partial \overline{x}^{j_1}} \cdots \frac{\partial x^{k_{r+1}}}{\partial \overline{x}^{j_{r+1}}},$$

which, if substituted in (2.17) establishes that $\delta^{i_1 \cdots i_{r+1}}_{j_1 \cdots j_{r+1}} \partial T_{i_1 \cdots i_r} / \partial x^{i_{r+1}}$ *is a* $(0, r + 1)$ *tensor.* We will return to this result in Section 5.2.

We return once more to (2.1) in order to introduce two additional numerical tensors, the *permutation symbols* (which are sometimes referred to as the

Levi–Civita symbols or alternating symbols). In (2.1) we set $r = n$ and define the permutation symbols $\varepsilon_{h_1 \cdots h_n}$, $\varepsilon^{j_1 \cdots j_n}$ by

$$\varepsilon_{h_1 \cdots h_n} = \delta^{1 2 \cdots n}_{h_1 \cdots h_n},$$

and $$\varepsilon^{j_1 \cdots j_n} = \delta^{j_1 \cdots j_n}_{1 2 \cdots n}, \tag{2.18}$$

respectively. We draw attention to the fact that each of these quantities has *exactly* $n$ indices, where $n$ is the dimension of $X_n$. From (2.18) and (2.2), $\varepsilon_{h_1 \cdots h_n}$ and $\varepsilon^{j_1 \cdots j_n}$ are each skew-symmetric in all their indices and will thus vanish if any of the indices coincide. Furthermore, from (2.18) and (2.1)

$$\varepsilon_{1 2 \cdots n} = \delta^{1 2 \cdots n}_{1 2 \cdots n} = 1,$$
$$\varepsilon^{1 2 \cdots n} = \delta^{1 2 \cdots n}_{1 2 \cdots n} = 1,$$

so that we have

$$\varepsilon_{h_1 \cdots h_n} = \varepsilon^{h_1 \cdots h_n} = \begin{cases} +1 & \text{if } h_1 \cdots h_n \text{ is an even permutation of } 1, 2, \ldots, n; \\ -1 & \text{if } h_1 \cdots h_n \text{ is an odd permutation of } 1, 2, \ldots, n; \\ 0 & \text{otherwise.} \end{cases}$$
$$\tag{2.19}$$

The terminology permutation symbol is thus justified.

We now wish to establish that

$$\varepsilon_{h_1 \cdots h_n} \varepsilon^{j_1 \cdots j_n} = \delta^{j_1 \cdots j_n}_{h_1 \cdots h_n}. \tag{2.20}$$

To this end consider the quantity

$$A^{j_1 \cdots j_n}_{h_1 \cdots h_n} \equiv \varepsilon_{h_1 \cdots h_n} \varepsilon^{j_1 \cdots j_n} - \delta^{j_1 \cdots j_n}_{h_1 \cdots h_n}, \tag{2.21}$$

which is clearly skew-symmetric in both sets of indices. Consequently, the only possible nonzero components of $A^{j_1 \cdots j_n}_{h_1 \cdots h_n}$ will occur when $j_1 \cdots j_n$ and $h_1 \cdots h_n$ are each a permutation of $1, 2, \ldots, n$. However, from (2.19), (2.21) and (2.1) it is easily seen that

$$A^{1 2 \cdots n}_{1 2 \cdots n} = 0.$$

We have thus shown that

$$A^{j_1 \cdots j_n}_{h_1 \cdots h_n} = 0,$$

which, by (2.21) establishes (2.20).

From (2.20) and (2.13) we also have

$$\varepsilon_{h_1 \cdots h_n} \varepsilon^{h_1 \cdots h_n} = n!. \tag{2.22}$$

If we consider the determinant of an arbitrary set of $n^2$ quantities $a_{ij}$ expressed in the form (1.3.2), we see that the summation taken together

with the factor $(-1)^\kappa$ is very closely related to (2.19). In fact (1.3.2) can be expressed in the form

$$\det(a_{ij}) = \varepsilon^{j_1 j_2 \cdots j_n} a_{1 j_1} a_{2 j_2} \cdots a_{n j_n}.$$

By virtue of (2.19) it is easily seen that this last expression can be written in the equivalent form

$$\varepsilon_{i_1 \cdots i_n} \det(a_{ij}) = \varepsilon^{j_1 \cdots j_n} a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}. \qquad (2.23)$$

In a similar way, if $b_j^i$ and $c^{ij}$ are each sets of $n^2$ quantities we have

$$\varepsilon_{i_1 \cdots i_n} \det(b_j^i) = \varepsilon_{j_1 \cdots j_n} b_{i_1}^{j_1} \cdots b_{i_n}^{j_n}, \qquad (2.24)$$

$$\varepsilon^{i_1 \cdots i_n} \det(b_j^i) = \varepsilon^{j_1 \cdots j_n} b_{j_1}^{i_1} \cdots b_{j_n}^{i_n}, \qquad (2.25)$$

and

$$\varepsilon^{i_1 \cdots i_n} \det(c^{ij}) = \varepsilon_{j_1 \cdots j_n} c^{i_1 j_1} \cdots c^{i_n j_n}. \qquad (2.26)$$

If we multiply (2.24) by $\varepsilon^{i_1 \cdots i_n}$, noting (2.22) and (2.20), we find

$$\det(b_j^i) = \frac{1}{n!} \delta_{j_1 \cdots j_n}^{i_1 \cdots i_n} b_{i_1}^{j_1} \cdots b_{i_n}^{j_n}, \qquad (2.27)$$

which is an expression for the determinant of $(b_j^i)$ in terms of the generalized Kronecker delta.

We now wish to determine the tensor character of the permutation symbols. In a new $\bar{x}$-coordinate system (2.18) becomes

$$\bar{\varepsilon}^{j_1 j_2 \cdots j_n} = \bar{\delta}_{1 2 \cdots n}^{j_1 \cdots j_n} = \frac{\partial \bar{x}^{j_1}}{\partial x^{k_1}} \frac{\partial \bar{x}^{j_2}}{\partial x^{k_2}} \cdots \frac{\partial \bar{x}^{j_n}}{\partial x^{k_n}} \frac{\partial x^{h_1}}{\partial \bar{x}^1} \frac{\partial x^{h_2}}{\partial \bar{x}^2} \cdots \frac{\partial x^{h_n}}{\partial \bar{x}^n} \delta_{h_1 \cdots h_n}^{k_1 \cdots k_n}, \qquad (2.28)$$

in which we substitute on the right-hand side from (2.20), noting at the same time that, by virtue of (2.24)

$$\varepsilon_{h_1 h_2 \cdots h_n} \frac{\partial x^{h_1}}{\partial \bar{x}^1} \frac{\partial x^{h_2}}{\partial \bar{x}^2} \cdots \frac{\partial x^{h_n}}{\partial \bar{x}^n} = J \qquad (2.29)$$

where $J$ denotes the Jacobian (1.5). Thus (2.28) reduces to

$$\bar{\varepsilon}^{j_1 \cdots j_n} = J \frac{\partial \bar{x}^{j_1}}{\partial x^{k_1}} \cdots \frac{\partial \bar{x}^{j_n}}{\partial x^{k_n}} \varepsilon^{k_1 \cdots k_n}, \qquad (2.30)$$

which demonstrates that *the permutation symbols $\varepsilon^{k_1 \cdots k_n}$ constitute the components of a type $(n, 0)$ relative tensor of weight $+1$, that is, a type $(n, 0)$ tensor density.*

Similarly, it may be shown that *the permutation symbol $\varepsilon_{h_1 \cdots h_n}$ constitute the components of a type $(0, n)$ relative tensor of weight $-1$.*

We shall now derive some important results concerning determinants. Again, let us consider an $(n \times n)$ matrix $(a_j^h)$, with $a = \det(a_j^h)$. The sum of all $(1 \times 1)$ principal minors is simply the trace, denoted by $a_{(1)}$:

$$a_{(1)} = a_j^j = \delta_h^j a_j^h. \qquad (2.31)$$

The sum of all $(2 \times 2)$ principal minors is, by definition,

$$a_{(2)} = (a_1^1 a_2^2 - a_1^2 a_2^1) + (a_1^1 a_3^3 - a_3^1 a_1^3) + \cdots + (a_{n-1}^{n-1} a_n^n - a_n^{n-1} a_{n-1}^n)$$
$$= \delta_{hk}^{12} a_1^h a_2^k + \delta_{hk}^{13} a_1^h a_3^k + \cdots + \delta_k^{n-1 \, n} a_{n-1}^h a_n^k.$$

In applying the summation convention to this sum, it should be noted that

$$\delta_{hk}^{lm} a_l^h a_m^k = \delta_{hk}^{12} a_1^h a_2^k + \delta_{hk}^{21} a_2^h a_1^k + \cdots = \delta_{hk}^{12} a_1^h a_2^k - \delta_{hk}^{12} a_2^h a_1^k + \cdots$$
$$= \delta_{hk}^{12} a_1^h a_2^k + \delta_{kh}^{12} a_1^k a_2^h + \cdots = 2(\delta_{hk}^{12} a_1^h a_2^k + \cdots),$$

so that we may write

$$2! \, a_{(2)} = \delta_{hk}^{lm} a_l^h a_m^k. \qquad (2.32)$$

Similarly, the sum $a_{(3)}$ of all $(3 \times 3)$ principal minors is seen to be given by

$$3! \, a_{(3)} = \delta_{hkq}^{lmp} a_l^h a_m^k a_p^q. \qquad (2.33)$$

Proceeding in this manner the following important formula is established for *the sum $a_{(s)}$ of all $(s \times s)$ principal minors of $\det(a_j^h)$*:

$$s! \, a_{(s)} = \delta_{h_1 h_2 \cdots h_s}^{j_1 j_2 \cdots j_s} a_{j_1}^{h_1} a_{j_2}^{h_2} \cdots a_{j_s}^{h_s} \qquad (s = 1, \ldots, n). \qquad (2.34)$$

We shall apply this result to the theory of *characteristic* (or *secular*) *determinants*, which play a basic role in linear algebra and in the theory of quadratic forms. Let us put, for the given matrix,

$$c_j^h = a_j^h + \lambda \delta_j^h, \qquad (2.35)$$

where $\lambda$ denotes an arbitrary parameter. The so-called characteristic determinant is, by definition,

$$c = \det(a_j^h + \lambda \delta_j^h), \qquad (2.36)$$

which is a polynomial of the $n$th order in $\lambda$. One is frequently interested in the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the equation $c = 0$, and accordingly it is useful to have explicit expressions for the coefficients of $\lambda, \lambda^2, \ldots, \lambda^n$ in the expansion of the characteristic determinant (2.36). These expressions are determined easily as follows. Using (2.35) and (2.27), we have

$$n! \, c = \delta_{h_1 h_2 \cdots h_n}^{j_1 j_2 \cdots j_n} (a_{j_1}^{h_1} + \lambda \delta_{j_1}^{h_1})(a_{j_2}^{h_2} + \lambda \delta_{j_2}^{h_2}) \cdots (a_{j_n}^{h_n} + \lambda \delta_{j_n}^{h_n}). \qquad (2.37)$$

Clearly the coefficient of $\lambda^n$ is $n!$, while the coefficient of $\lambda^{n-s}$ (for $s \geq 1$) consists of the sum of $\binom{n}{s}$ terms of the type

$$\delta_{h_1 \cdots h_s h_{s+1} \cdots h_n}^{j_1 \cdots j_s j_{s+1} \cdots j_n} a_{j_1}^{h_1} \cdots a_{j_s}^{h_s} \delta_{j_{s+1}}^{h_{s+1}} \cdots \delta_{j_n}^{h_n} = \delta_{h_1 \cdots h_s h_{s+1} \cdots h_n}^{j_1 \cdots j_s j_{s+1} \cdots j_n} a_{j_1}^{h_1} \cdots a_{j_s}^{h_s}$$
$$= (n-s)! \, \delta_{h_1 \cdots h_s}^{j_1 \cdots j_s} a_{j_1}^{h_1} \cdots a_{j_s}^{h_s},$$

where, in the second step, we have used (2.11) with $r = n$. Since

$$\binom{n}{s} = \frac{n!}{(n-s)! \, s!},$$

it follows from (2.37) that the coefficient of $\lambda^{n-s}$ in $c$ is simply

$$\frac{1}{s!} \delta_{h_1 \cdots h_s}^{j_1 \cdots j_s} a_{j_1}^{h_1} \cdots a_{j_s}^{h_s}.$$

*Thus the characteristic determinant* (2.36) *may be expressed in the form*

$$\det(a_j^h + \lambda \delta_j^h) = \lambda^n + \sum_{s=1}^{n} \frac{\lambda^{n-s}}{s!} \delta_{h_1 \cdots h_s}^{j_1 \cdots j_s} a_{j_1}^{h_1} \cdots a_{j_s}^{h_s}. \tag{2.38}$$

By means of (2.34) this can be written more elegantly as

$$\det(a_j^h + \lambda \delta_j^h) = \lambda^n + \sum_{s=1}^{n} \lambda^{n-s} a_{(s)}, \tag{2.39}$$

in which $a_{(s)}$ again denotes the sum of all $s \times s$ principal minors of $\det(a_j^h)$.

The formulae (2.38) and (2.39) have an immediate corollary which will be required later. Let us denote the $s$th elementary symmetric function of the roots $\lambda_1, \ldots, \lambda_n$ of the characteristic equation

$$\det(a_j^h + \lambda \delta_j^h) = 0 \tag{2.40}$$

by $H_{(s)}$, so that

$$\left.\begin{array}{l} H_{(1)} = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \\ H_{(2)} = \lambda_1 \lambda_2 + \lambda_1 \lambda_3 + \cdots + \lambda_{n-1} \lambda_n, \\ H_{(3)} = \lambda_1 \lambda_2 \lambda_3 + \lambda_1 \lambda_2 \lambda_4 + \cdots + \lambda_{n-2} \lambda_{n-1} \lambda_n, \\ H_{(n)} = \lambda_1 \lambda_2 \lambda_3 \cdots \lambda_n. \end{array}\right\} \tag{2.41}$$

Then it is known from the elementary theory of polynomial equations that these quantities are related to the coefficients $a_{(s)}$ in (2.39) according to

$$a_{(s)} = (-1)^s H_{(s)}. \tag{2.42}$$

Because of (2.34) we thus obtain *an explicit expression for the* $s$th *elementary symmetric function of the roots of* (2.40), *namely,*

$$H_{(s)} = (-1)^s \frac{1}{s!} \delta_{h_1 h_2 \cdots h_s}^{j_1 j_2 \cdots j_s} a_{j_1}^{h_1} a_{j_2}^{h_2} \cdots a_{j_s}^{h_s}. \tag{2.43}$$

Finally, we shall obtain a useful expression for the cofactors of the elements $a_j^h$ in the determinant $\det(a_j^h)$. To this end let us consider the quantities $A_h^j$ defined by

$$(n-1)! \, A_h^j = \delta_{hh_2 \cdots h_n}^{jj_2 \cdots j_n} a_{j_2}^{h_2} \cdots a_{j_n}^{h_n}. \tag{2.44}$$

Let us multiply (2.44) by $a_l^h$, after which we apply (2.20), which gives

$$(n-1)! \, A_h^j a_l^h = \varepsilon^{jj_2 \cdots j_n} \varepsilon_{hh_2 \cdots h_n} a_l^h a_{j_2}^{h_2} \cdots a_{j_n}^{h_n}.$$

With the aid of (2.20) and (2.24) this becomes

$$(n-1)! \, A_h^j a_l^h = \varepsilon^{jj_2 \cdots j_n} a \varepsilon_{lj_2 \cdots j_n} = a \delta_{lj_2 \cdots j_n}^{jj_2 \cdots j_n},$$

or, because of (2.11), with $s = 1$ and $r = n$,

$$A_h^j a_l^h = a \delta_l^j. \tag{2.45}$$

*Thus the quantities* $A_h^j$ *as defined in* (2.44) *are the cofactors of the elements* $a_j^h$ *in the determinant* $\det(a_j^h)$.

## 4.3  NORMAL COORDINATES

In Section 3.7 we encountered a class of curves which are characterized by the fact that successive tangents result from each other by parallel displacement. These so-called autoparallel curves or paths satisfy the system of $n$ second-order differential equations (3.7.17), namely,

$$\frac{d^2 x^j}{dt^2} + \Gamma_{h \ k}^{\ j} \frac{dx^h}{dt} \frac{dx^k}{dt} = 0, \tag{3.1}$$

in which the $\Gamma_{h \ k}^{\ j}$ denote the coefficients of a given connection on our differentiable manifold $X_n$, while $t$ is a prescribed parameter.

We shall now show that a pair of points $O$ and $P$ of $X_n$ can be joined by a unique path, provided that $O$ and $P$ are sufficiently close to each other. In the course of our derivation of this assertion we shall construct a special coordinate system, whose properties will be studied in some detail.

Let us differentiate (3.1) with respect to $t$:

$$\frac{d^3 x^j}{dt^3} + \frac{\partial \Gamma_{h \ k}^{\ j}}{\partial x^l} \frac{dx^h}{dt} \frac{dx^k}{dt} \frac{dx^l}{dt} + \Gamma_{h \ k}^{\ j} \frac{d^2 x^h}{dt^2} \frac{dx^k}{dt} + \Gamma_{h \ k}^{\ j} \frac{dx^h}{dt} \frac{d^2 x^k}{dt^2} = 0,$$

or, if we replace the second derivatives of $x^h$ and $x^k$ according to (3.1),

$$\frac{d^3 x^j}{dt^3} + \Gamma_{h \ kl}^{\ j} \frac{dx^h}{dt} \frac{dx^k}{dt} \frac{dx^l}{dt} = 0, \tag{3.2}$$

in which we have put, for the sake of brevity,

$$\Gamma_{h\ kl}^{\ j} = \frac{\partial \Gamma_{h\ k}^{\ j}}{\partial x^l} - \Gamma_{m\ k}^{\ j}\Gamma_{l\ h}^{\ m} - \Gamma_{h\ m}^{\ j}\Gamma_{k\ l}^{\ m}. \tag{3.3}$$

We note that (3.2) yields the third derivatives $d^3x^j/dt^3$ as polynomials of the third order in $dx^k/dt$; similarly, by differentiation of (3.2) we find that the fourth derivatives $d^4x^j/dt^4$ can be expressed as polynomials of the fourth order in $dx^k/dt$. Clearly this process can be repeated indefinitely, provided that the connection coefficients $\Gamma_{h\ k}^{\ j}$ are of class $C^\infty$, as is assumed here. The precise form of the coefficients of these polynomials, such as (3.3), does not concern us in this context, and accordingly the corresponding expressions will not be derived.

Now let us consider a fixed point $O$ of $X_n$, and a path $\Gamma$, which passes through $O$ with given tangent vector $\dot{x}^h = dx^h/dt$ at $O$. It is assumed that the parameter $t$ is chosen such that $t = 0$ at $O$. Let $x^j(t)$ denote the coordinates of a point $P$ on $\Gamma$ corresponding to the parameter value $t$. By means of a Taylor series expansion we can express the coordinates $x^j(t)$ of $P$ as follows:

$$x^j(t) = x^j(0) + t\left(\frac{dx^j}{dt}\right)_{t=0} + \frac{t^2}{2!}\left(\frac{d^2x^j}{dt^2}\right)_{t=0} + \frac{t^3}{3!}\left(\frac{d^3x^j}{dt^3}\right)_{t=0} + \cdots,$$

or, if we use (3.1), (3.2), and the relations obtained by subsequent differentiation of (3.2),

$$x^j(t) = x^j(0) + t\dot{x}^j - \frac{t^2}{2!}\Gamma_{h\ k}^{\ j}\dot{x}^h\dot{x}^k - \frac{t^3}{3!}\Gamma_{h\ kl}^{\ j}\dot{x}^h\dot{x}^k\dot{x}^l - \cdots, \tag{3.4}$$

where it is to be understood that $\Gamma_{h\ k}^{\ j}, \Gamma_{h\ kl}^{\ j}, \ldots$, are evaluated at $O$, while $\dot{x}^j$ is the tangent vector of $\Gamma$ at $O$.

Clearly this series is convergent for sufficiently small values of $t$; we therefore suppose that the point $P$ has been chosen such that this is indeed the case. Thus, at least locally, the path $\Gamma$ is completely specified by the initial point $O$ and the initial direction $\dot{x}^h$.

Now let us introduce a new set of quantities $\xi^j$ by writing

$$\xi^j = t\dot{x}^j. \tag{3.5}$$

Then (3.4) may be written in the form

$$x^j(t) = x^j(0) + \xi^j - \frac{1}{2!}\Gamma_{h\ k}^{\ j}\xi^h\xi^k - \frac{1}{3!}\Gamma_{h\ kl}^{\ j}\xi^h\xi^k\xi^l - \cdots. \tag{3.6}$$

Differentiation of (3.6) with respect to $\xi^p$ yields

$$\frac{\partial x^j}{\partial \xi^p} = \delta_p^j - \frac{1}{2!}(\Gamma_{p\ k}^{\ j}\xi^k + \Gamma_{h\ p}^{\ j}\xi^h)$$

$$- \frac{1}{3!}(\Gamma_{p\ kl}^{\ j}\xi^k\xi^l + \Gamma_{h\ pl}^{\ j}\xi^h\xi^l + \Gamma_{h\ kp}^{\ j}\xi^h\xi^k) - \cdots. \tag{3.7}$$

From (3.5) it is obvious that $\xi^j \to 0$ when $t \to 0$. Therefore, at $O$,

$$\frac{\partial x^j}{\partial \xi^p} = \delta_p^j, \tag{3.8}$$

and accordingly the Jacobian

$$\frac{\partial(x^1, \ldots, x^n)}{\partial(\xi^1, \ldots, \xi^n)} = 1 \tag{3.9}$$

at $O$.

We shall regard the variables $\xi^j$ as new coordinates, the corresponding coordinate transformation

$$x^j = x^j(\xi^h) \tag{3.10}$$

being given explicitly by (3.6); because of (3.9) this transformation possesses an inverse

$$\xi^h = \xi^h(x^j) \tag{3.11}$$

in a finite neighborhood $U$ of $O$.

Thus, given a point $P$ with coordinates $x^j$ in the neighborhood $U$ of $O$, its coordinates $\xi^h$ in the new coordinate system are uniquely determined by virtue of (3.11). Because of (3.5), this defines the quantities $t\dot{x}^j$ at $O$, which in turn yield the path through $O$ and $P$ in view of (3.4). We have therefore proved that, *for any point $O$ of our differentiable manifold $X_n$, there exists a neighborhood $U$ of $O$ such that any point $P \in U$ is joined to $O$ by a unique path.* Furthermore, it is easily seen that this path lies entirely in $U$.

Let us evaluate the connection coefficients $\bar{\Gamma}_{p\ q}^{\ s}$ in the $\xi^j$-coordinate system at $O$. This is done by means of the transformation law (3.3.22), which, in this context, assumes the form

$$\frac{\partial x^j}{\partial \xi^s}\bar{\Gamma}_{p\ q}^{\ s} = \Gamma_{h\ k}^{\ j}\frac{\partial x^h}{\partial \xi^p}\frac{\partial x^k}{\partial \xi^q} + \frac{\partial^2 x^j}{\partial \xi^p\,\partial \xi^q}. \tag{3.12}$$

The values of the first derivatives $\partial \xi^s/\partial x^j$, $\partial x^h/\partial \xi^p$ at $O$ are known in terms of (3.8). The second derivatives are calculated by differentiation of (3.7) with respect to $\xi^q$, which gives

$$\frac{\partial^2 x^j}{\partial \xi^p\,\partial \xi^q} = -\tfrac{1}{2}(\Gamma_{p\ q}^{\ j} + \Gamma_{q\ p}^{\ j}) + \cdots,$$

in which the dots $\cdots$ denote a polynomial in $\xi^j$, beginning with first-order terms. Again, because $\xi^j = 0$ at $O$, we infer that

$$\frac{\partial^2 x^j}{\partial \xi^p \partial \xi^q} = -\tfrac{1}{2}(\Gamma_p{}^j{}_q + \Gamma_q{}^j{}_p) \qquad (3.13)$$

at $O$. We now substitute (3.8) and (3.13) in (3.12), which yields

$$\delta_s^j \overline{\Gamma}_p{}^s{}_q = \Gamma_h{}^j{}_k \delta_p^h \delta_q^k - \tfrac{1}{2}(\Gamma_p{}^j{}_q + \Gamma_q{}^j{}_p),$$

or

$$\overline{\Gamma}_p{}^j{}_q = \Gamma_p{}^j{}_q - \tfrac{1}{2}(\Gamma_p{}^j{}_q + \Gamma_q{}^j{}_p) = \tfrac{1}{2}(\Gamma_p{}^j{}_q - \Gamma_q{}^j{}_p) \qquad (3.14)$$

at $O$.

Thus the transformed connection coefficients $\overline{\Gamma}_p{}^j{}_q$ at $O$ are identical with the skew-symmetric part of the given connection.

In particular, *if the given connection is symmetric, the coefficients of the transformed connection vanish at $O$.* Under these circumstances the coordinate system defined by (3.11) is called a *normal coordinate system*, and the point $O$ at which the connection coefficients are zero in this system is known as the *pole* of these coordinates.

It has therefore been established that, given any *symmetric* class $C^\infty$ connection on the differentiable manifold $X_n$, *one can always construct a coordinate system whose pole is located at an arbitrary point $O$ such that the transformed connection coefficients vanish at $O$.*

*Remark 1.* The condition that the given connection be symmetric is essential. This is not only obvious from (3.14), but also is a direct result of the fact that the skew-symmetric part $S_h{}^j{}_k$ of any nonsymmetric connection is tensorial: if $S_h{}^j{}_k$ does not vanish at $O$ in the given coordinate system, it cannot vanish at $O$ in any other system, and therefore there does not exist a coordinate system in which the nonsymmetric connection coefficients are zero at $O$.

*Remark 2.* It should be clearly understood that, in the case of a symmetric connection, the transformed connection coefficients vanish solely at the pole $O$ of the normal coordinates. The first- and higher-order derivatives of the transformed coefficients do not generally vanish at $O$. This fact must always be borne in mind when normal coordinates are used. In general there do *not* exist coordinate systems on $X_n$ in which the connection coefficients vanish on a finite region of $X_n$. Indeed, under such conditions it would follow from the definition (3.6.8) that the curvature tensor would vanish identically on that region, which again is a tensorial condition that cannot generally be satisfied.

*Remark 3.* From the definition (3.5) it is evident that relative to our special coordinate system the coordinates of any point $P$ on the path $\Gamma$ through the pole $O$ are given by $\xi^j = \dot{x}^j t$, in which $t$ is variable, while $\dot{x}^j$ is fixed. Thus the differential equations satisfied by $\Gamma$ are simply

$$\frac{d^2 \xi^j}{dt^2} = 0. \qquad (3.15)$$

If this is compared with the general system (3.1) relative to the $\xi^h$-coordinates, it follows that

$$\overline{\Gamma}_h{}^j{}_k \frac{d\xi^h}{dt} \frac{d\xi^k}{dt} = 0 \qquad (3.16)$$

in the special coordinate system along any path through the pole $O$. However, since the coefficients of $\overline{\Gamma}_h{}^j{}_k$ in (3.16) are not arbitrary, this does not imply that $\overline{\Gamma}_h{}^j{}_k$ vanishes along these paths. Because of (3.5) we can regard our special coordinates as a generalization of the concept of spherical polar coordinates of Euclidean geometry with pole or origin at $O$.

*Remark 4.* Normal coordinates can sometimes be used to simplify certain calculations whenever a *symmetric* connection is given. For instance, at the pole $O$ of a normal coordinate system the components $\overline{K}_l{}^j{}_{hk}$ of the curvature tensor (3.6.8) reduce to

$$\overline{K}_l{}^j{}_{hk} = \frac{\partial \overline{\Gamma}_l{}^j{}_h}{\partial \xi^k} - \frac{\partial \overline{\Gamma}_l{}^j{}_k}{\partial \xi^h},$$

from which the identity (3.8.7) follows directly. Since this identity involves only tensors its validity in all coordinate systems is thus established. Furthermore, at the pole of a normal coordinate system the covariant derivative of a tensor or tensor density is identical with its usual derivative. By means of this fact it is a simple matter to derive the Bianchi identities (3.8.12) by direct differentiation of the expression (3.6.8) for the curvature tensor at $O$ in normal coordinates.

## 4.4   THE LIE DERIVATIVE

We saw in Section 3.3 that if our differentiable manifold is endowed with an affine connection, a process of tensorial differentiation of tensor fields can be introduced. In contrast to this we now consider the situation when our manifold is endowed with a type $(1, 0)$ tensor field, the components of which will be denoted by $v^i$. By using arguments quite similar to those presented in Section 3.4, we show that it is again possible to introduce a tensorial process of differentiation, which, although approached in the same spirit, differs radically from covariant differentiation.

For the sake of simplicity we illustrate the process for the case of a type $(1, 1)$ tensor density field, the components $T_h^p$ of which transform according to

$$\bar{T}_l^j(\bar{x}^k) = J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} T_h^p(x^k), \tag{4.1}$$

where $J$ is defined by (1.5). The partial derivatives relative to the two coordinate systems are related by

$$\frac{\partial \bar{T}_l^j}{\partial \bar{x}^k} = J \frac{\partial^2 \bar{x}^j}{\partial x^s \partial x^p} \frac{\partial x^s}{\partial \bar{x}^k} \frac{\partial x^h}{\partial \bar{x}^l} T_h^p + J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial^2 x^h}{\partial \bar{x}^k \partial \bar{x}^l} T_h^p$$
$$+ J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} \frac{\partial x^s}{\partial \bar{x}^k} \frac{\partial T_h^p}{\partial x^s} + \frac{\partial J}{\partial \bar{x}^k} \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} T_h^p. \tag{4.2}$$

We also have

$$\bar{v}^k = \frac{\partial \bar{x}^k}{\partial x^q} v^q. \tag{4.3}$$

Thus a two-index quantity is obtained from (4.2) by multiplication of (4.2) by $\bar{v}^k$:

$$\frac{\partial \bar{T}_l^j}{\partial \bar{x}^k} \bar{v}^k = J \frac{\partial^2 \bar{x}^j}{\partial x^s \partial x^p} v^s \frac{\partial x^h}{\partial \bar{x}^l} T_h^p + J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial^2 x^h}{\partial \bar{x}^k \partial \bar{x}^l} \bar{v}^k T_h^p$$
$$+ J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} \frac{\partial T_h^p}{\partial x^s} v^s + \frac{\partial J}{\partial \bar{x}^k} \bar{v}^k \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} T_h^p. \tag{4.4}$$

Clearly the first, second, and fourth terms on the right-hand side of (4.4) are the ones to be rewritten since they contain second-order derivatives [see (1.19)]. From (4.3) we have, by differentiation with respect to $x^p$,

$$\frac{\partial \bar{x}^k}{\partial x^p} \frac{\partial \bar{v}^j}{\partial \bar{x}^k} = \frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^s} v^s + \frac{\partial \bar{x}^j}{\partial x^s} \frac{\partial v^s}{\partial x^p},$$

so that the first term on the right-hand side of (4.4) can be written as

$$J \frac{\partial^2 \bar{x}^j}{\partial x^p \partial x^s} v^s \frac{\partial x^h}{\partial \bar{x}^l} T_h^p = \frac{\partial \bar{v}^j}{\partial \bar{x}^k} \bar{T}_l^k - J \frac{\partial \bar{x}^j}{\partial x^s} \frac{\partial x^h}{\partial \bar{x}^l} \frac{\partial v^s}{\partial x^p} T_h^p, \tag{4.5}$$

where use has been made of (4.1). From (4.3) we also have

$$\frac{\partial x^h}{\partial \bar{x}^k} \bar{v}^k = v^h,$$

which, when differentiated with respect to $\bar{x}^l$, yields

$$\frac{\partial^2 x^h}{\partial \bar{x}^l \partial \bar{x}^k} \bar{v}^k = - \frac{\partial x^h}{\partial \bar{x}^k} \frac{\partial \bar{v}^k}{\partial \bar{x}^l} + \frac{\partial x^t}{\partial \bar{x}^l} \frac{\partial v^h}{\partial x^t}. \tag{4.6}$$

This allows the second term on the right-hand side of (4.4) to be expressed in the form

$$J \frac{\partial^2 x^h}{\partial \bar{x}^l \partial \bar{x}^k} \bar{v}^k \frac{\partial \bar{x}^j}{\partial x^p} T_h^p = - \frac{\partial \bar{v}^k}{\partial \bar{x}^l} \bar{T}_k^j + J \frac{\partial x^k}{\partial \bar{x}^l} \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial v^h}{\partial x^k} T_h^p. \tag{4.7}$$

From (1.19) we have

$$\frac{\partial J}{\partial \bar{x}^k} \bar{v}^k = J \frac{\partial^2 x^h}{\partial \bar{x}^l \partial \bar{x}^k} \bar{v}^k \frac{\partial \bar{x}^l}{\partial x^h},$$

which, together with (4.6), enables one to write the fourth term on the right-hand side of (4.4) as follows:

$$\frac{\partial J}{\partial \bar{x}^k} \bar{v}^k \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} T_h^p = - \frac{\partial \bar{v}^k}{\partial \bar{x}^k} \bar{T}_l^j + J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} \frac{\partial v^k}{\partial x^k} T_h^p. \tag{4.8}$$

We now substitute (4.5), (4.7), and (4.8) in (4.4) to find

$$\frac{\partial \bar{T}_l^j}{\partial \bar{x}^k} \bar{v}^k - \frac{\partial \bar{v}^j}{\partial \bar{x}^k} \bar{T}_l^k + \frac{\partial \bar{v}^k}{\partial \bar{x}^l} \bar{T}_k^j + \frac{\partial \bar{v}^k}{\partial \bar{x}^k} \bar{T}_l^j$$
$$= J \frac{\partial \bar{x}^j}{\partial x^p} \frac{\partial x^h}{\partial \bar{x}^l} \left[ \frac{\partial T_h^p}{\partial x^s} v^s - \frac{\partial v^p}{\partial x^k} T_h^k + \frac{\partial v^k}{\partial x^h} T_k^p + \frac{\partial v^k}{\partial x^k} T_h^p \right]. \tag{4.9}$$

This suggests that we define the so-called *Lie derivative* of $T_h^p$ as follows:

$$\text{£}_v T_h^p = \frac{\partial T_h^p}{\partial x^s} v^s - T_h^k \frac{\partial v^p}{\partial x^k} + T_k^p \frac{\partial v^k}{\partial x^h} + T_h^p \frac{\partial v^k}{\partial x^k}, \tag{4.10}$$

for (4.9) clearly indicates that these quantities are the components of a type $(1, 1)$ tensor density.

This same process, applied to the general transformation law (1.14), yields the following expression for *the Lie derivative*† *of a type* $(r, s)$ *relative tensor field of weight* $w$:

$$\text{£}_v \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_s} = \frac{\partial \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}}{\partial x^h} v^h - \sum_{\alpha=1}^{r} \Lambda^{j_1 \cdots j_{\alpha-1} m j_{\alpha+1} \cdots j_r}{}_{l_1 \cdots l_s} \frac{\partial v^{j_\alpha}}{\partial x^m}$$
$$+ \sum_{\beta=1}^{s} \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_{\beta-1} m l_{\beta+1} \cdots l_s} \frac{\partial v^m}{\partial x^{l_\beta}} + w \frac{\partial v^k}{\partial x^k} \Lambda^{j_1 \cdots j_r}{}_{l_1 \cdots l_s}. \tag{4.11}$$

The *basic laws of Lie differentiation* are very similar to the corresponding laws for absolute differentiation:

1. The Lie derivative of a type $(r, s)$ relative tensor field is a type $(r, s)$ relative tensor of the same weight.

† For an extensive study of the Lie derivative see Yano [1].

2.   The Lie derivative of the sum of two relative tensor fields of the same type and weight is the sum of the Lie derivatives of these fields.

3.   The product rule for Lie differentiation is formally identical with the product rule of ordinary differentiation.

Statement 1 is valid as a result of the manner in which the Lie derivative is constructed. Statement 2 follows from the linearity of (4.11) in the tensor field. It is possible to establish statement 3 in complete generality. However, in order to avoid a profusion of indices we will content ourselves with establishing it for the special case of the product of a type $(2, 0)$ tensor field with a type $(0, 1)$ tensor density field. Consequently, we shall consider the product

$$T^{ij}{}_h = U^{ij}V_h. \tag{4.12}$$

From (4.11) we have

$$\pounds_v T^{ij}{}_h = \frac{\partial T^{ij}{}_h}{\partial x^k} v^k - T^{pj}{}_h \frac{\partial v^i}{\partial x^p} - T^{ip}{}_h \frac{\partial v^j}{\partial x^p} + T^{ij}{}_p \frac{\partial v^p}{\partial x^h} + T^{ij}{}_h \frac{\partial v^k}{\partial x^k},$$

which, by (4.12), yields

$$\pounds_v T^{ij}{}_h = \left[\frac{\partial U^{ij}}{\partial x^k} v^k - U^{pj} \frac{\partial v^i}{\partial x^p} - U^{ip} \frac{\partial v^j}{\partial x^p}\right]V_h + U^{ij}\left[\frac{\partial V_h}{\partial x^k} v^k + V_p \frac{\partial v^p}{\partial x^h} + V_h \frac{\partial v^k}{\partial x^k}\right].$$

Because of (4.11) the latter reduces to

$$\pounds_v(U^{ij}V_h) = (\pounds_v U^{ij})V_h + U^{ij}(\pounds_v V_h),$$

as required.

*Remark 1.*   If, in addition to the type $(1, 0)$ tensor field, our manifold is endowed with an affine connection, we can express (4.11) in terms of partial covariant derivatives as follows. From (3.5.4) and (3.6.7) we have

$$\frac{\partial v^j}{\partial x^m} = v^j{}_{|m} - \Gamma_{h}{}^{j}{}_{m} v^h = v^j{}_{|m} - S_h{}^{j}{}_m v^h - \Gamma_m{}^{j}{}_h v^h,$$

which, when substituted in (4.11), account being taken of (1.24), yields

$$\pounds_v \Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_s} = \Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_s|h} v^h - \sum_{\alpha=1}^{r} \Lambda^{j_1\cdots j_{\alpha-1}m j_{\alpha+1}\cdots j_r}{}_{l_1\cdots l_s}(v^{j_\alpha}{}_{|m} - S_h{}^{j_\alpha}{}_m v^h)$$
$$+ \sum_{\beta=1}^{s} \Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_{\beta-1}m l_{\beta+1}\cdots l_s}(v^m{}_{|l_\beta} - S_h{}^{m}{}_{l_\beta} v^h) + w v^k{}_{|k} \Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_s}. \tag{4.13}$$

In this form the tensorial character of the Lie derivative is obvious.

*Remark 2.*   It is possible to introduce the Lie derivative in an alternative way as follows (Yano [1]). With the given vector field we associate the trans-

formation

$$\bar{x}^i = x^i + \lambda v^i(x^j) \tag{4.14}$$

where $\lambda$ is an arbitrary constant. The Lie derivative of $\Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_s}(x^h)$ is then defined by

$$\pounds_v \Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_s} = \lim_{\lambda\to 0} \frac{\Lambda^{j_1\cdots j_r}{}_{l_1\cdots l_s}(x^h) - \overline{\Lambda}^{j_1\cdots j_r}{}_{l_1\cdots l_s}(x^h)}{\lambda}. \tag{4.15}$$

It is not difficult to show the equivalence of (4.11) and (4.15). However, again for the sake of simplicity, we shall content ourselves with establishing this equivalence only in the case (4.1).

From (4.14) we have

$$\frac{\partial \bar{x}^i}{\partial x^j} = \delta^i_j + \lambda \frac{\partial v^i}{\partial x^j} \tag{4.16}$$

from which we obtain

$$\det\left(\frac{\partial \bar{x}^i}{\partial x^j}\right) = \det\left(\delta^i_j + \lambda \frac{\partial v^i}{\partial x^j}\right) = \lambda^n \det\left(\frac{\partial v^i}{\partial x^j} + \lambda^{-1}\delta^i_j\right). \tag{4.17}$$

From (2.38) we thus see that from (4.17)

$$\det\left(\frac{\partial \bar{x}^i}{\partial x^j}\right) = 1 + \lambda \frac{\partial v^i}{\partial x^i} + \cdots \tag{4.18}$$

where $\cdots$ denotes terms at least of order $\lambda^2$. We can write (4.1) in the form

$$\det\left(\frac{\partial \bar{x}^i}{\partial x^s}\right)\frac{\partial \bar{x}^l}{\partial x^h} \overline{T}^j_l(\bar{x}^k) = \frac{\partial \bar{x}^j}{\partial x^p} T^p_h(x^k), \tag{4.19}$$

which, by (4.16) and (4.18), reduces to

$$\left(1 + \lambda \frac{\partial v^i}{\partial x^i} + \cdots\right)\left(\delta^l_h + \lambda \frac{\partial v^l}{\partial x^h}\right)\overline{T}^j_l(\bar{x}^k) = \left(\delta^j_p + \lambda \frac{\partial v^j}{\partial x^p}\right)T^p_h(x^k)$$

or

$$\overline{T}^j_h(\bar{x}^k) + \lambda\left[\frac{\partial v^i}{\partial x^i}\overline{T}^j_h(\bar{x}^k) + \frac{\partial v^l}{\partial x^h}\overline{T}^j_l(\bar{x}^k)\right] + \cdots = T^j_h(x^k) + \lambda \frac{\partial v^j}{\partial x^p}T^p_h(x^k). \tag{4.20}$$

However, by expanding $\overline{T}^j_l(\bar{x}^k)$ in a Taylor series about $x^i$ we have

$$\overline{T}^j_h(\bar{x}^k) = \overline{T}^j_h(x^k) + \left(\frac{\partial \overline{T}^j_h}{\partial \bar{x}^k}\right)_{\lambda=0}\lambda v^k + \cdots. \tag{4.21}$$

Furthermore, it is easily seen from (4.20) that

$$\left(\frac{\partial \overline{T}^j_h}{\partial \bar{x}^k}\right)_{\lambda=0} = \frac{\partial T^j_h}{\partial x^k},$$

so that (4.20) and (4.21) give

$$T_h^j(x^k) - \overline{T}_h^j(x^k) = \lambda\left[\frac{\partial T_h^j}{\partial x^k}v^k - T_h^p\frac{\partial v^j}{\partial x^p} + \overline{T}_l^j(\overline{x}^k)\frac{\partial v^l}{\partial x^h} + \frac{\partial v^i}{\partial x^i}\overline{T}_h^j(\overline{x}^k)\right] + \cdots.$$

We now divide this equation by $\lambda$ and let $\lambda \to 0$, noting that by (4.20)

$$\lim_{\lambda \to 0} \overline{T}_h^j(\overline{x}^k) = T_h^j(x^k),$$

to find

$$\lim_{\lambda \to 0}\frac{T_h^j(x^k) - \overline{T}_h^j(x^k)}{\lambda} = \frac{\partial T_h^j}{\partial x^k}v^k - T_h^p\frac{\partial v^j}{\partial x^p} + T_l^j\frac{\partial v^l}{\partial x^h} + \frac{\partial v^i}{\partial x^i}T_h^j,$$

which is in complete agreement with (4.10).

Originally the Lie derivative was defined in a form similar to that given by (4.15) which arose quite naturally in the theory of deformations of sub-spaces of a differentiable manifold (Davies [1]).

*Remark 3.* The above analysis (4.1)–(4.10) can be extended to certain quantities which are not relative tensors, although the associated Lie derivative may well be tensorial. For example, the Lie derivative of an affine connection is given by

$$\pounds_v\Gamma_{jk}^i = \frac{\partial\Gamma_{jk}^i}{\partial x^h}v^h + \frac{\partial^2 v^i}{\partial x^j\partial x^k} - \Gamma_{jk}^l\frac{\partial v^i}{\partial x^l} + \Gamma_{hk}^i\frac{\partial v^h}{\partial x^j} + \Gamma_{jh}^i\frac{\partial v^h}{\partial x^k},$$

which may be expressed in the equivalent form

$$\pounds_v\Gamma_{jk}^i = v^hK_{jkh}^i + (v^hS_{jh}^i)_{|k} + v^i_{|j|k},$$

which clearly demonstrates that these are the components of a type (1, 2) tensor.

## PROBLEMS

Unless otherwise stated, all quantities are the components of absolute tensors of the type indicated by the position of their indices.

4.1   If $T^{rs}$ are the components of a type (2, 0) tensor density show that

$$T^{rs}_{|r} = \frac{\partial T^{rs}}{\partial x^r} + \Gamma_r^s{}_r T^{rh}.$$

If the connection is symmetric and $T^{rs} = -T^{sr}$ show that

$$T^{rs}_{|r} = \frac{\partial T^{rs}}{\partial x^r}.$$

4.2   (a) Prove that the covariant derivative of the generalized Kronecker delta vanishes identically for any connection.
   (b) Prove that the covariant derivatives of the permutation symbols vanish identically, if the connection is symmetric.

4.3   Give an expression for the general $(3 \times 3)$ determinant formed from the matrix

$$\begin{pmatrix} A_1 & A_2 & \cdots & A_n \\ B_1 & B_2 & \cdots & B_n \\ C_1 & C_2 & \cdots & C_n \end{pmatrix}.$$

4.4   In $X_3$, with a symmetric connection show that

$$\varepsilon^{ijk}A_{j|i} = \varepsilon^{ijk}\frac{\partial A_j}{\partial x^i}$$

and

$$(\varepsilon^{ijk}A_jB_k)_{|i} = \varepsilon^{ijk}A_{j|i}B_k - \varepsilon^{ijk}B_{j|i}A_k.$$

[*Note*: this is the statement

$$\text{div}(\mathbf{A} \times \mathbf{B}) = (\text{curl }\mathbf{A})\cdot\mathbf{B} - (\text{curl }\mathbf{B})\cdot\mathbf{A}].$$

What are the statements (identities)

$$\text{curl grad } f = 0$$
$$\text{div curl } \mathbf{A} = 0$$

where $f$ is a scalar? Prove them using the index notation.

4.5   Relative to a rectangular coordinate system in $E_3$ Maxwell's equations may be written in the form

$$\text{curl }\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{H}}{\partial t}, \quad \text{div }\mathbf{E} = 4\pi\rho$$

$$\text{curl }\mathbf{H} = \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{J}, \quad \text{div }\mathbf{H} = 0.$$

Using the numerical tensors express these relations in component form valid in any coordinate system (Davis [1], Landau and Lifshitz [1], and Panofsky and Philips [1]).

4.6   If, in $X_4$, the connection is symmetric, show that

   (a) $\varepsilon^{ijkh}K_{jkh}^l = 0,$
   (b) $\varepsilon^{ijkh}K_{aij|k}^b = 0.$

*4.7   Prove that

$$\delta_{h_1\cdots h_sh_{s+1}\cdots h_r}^{j_1\cdots j_sj_{s+1}\cdots j_r}\delta_{k_{s+1}\cdots k_r}^{h_{s+1}\cdots h_r} = (r-s)!\,\delta_{h_1\cdots h_sk_{s+1}\cdots k_r}^{j_1\cdots j_sj_{s+1}\cdots j_r}.$$

4.8   Prove that

$$\varepsilon^{j_1\cdots j_rj_{r+1}\cdots j_n}\delta_{j_{r+1}\cdots j_n}^{h_{r+1}\cdots h_n} = (n-r)!\,\varepsilon^{j_1\cdots j_rh_{r+1}\cdots h_n}.$$

4.9  By considering $\delta^{ijklh}_{rstum} H^{um}{}_{lh}$, where $H^{um}{}_{lh}$ satisfies (2.6), show that, in $X_4$,

$$\delta^i_r H^{jk}{}_{st} + \delta^i_t H^{jk}{}_{rs} + \delta^i_s H^{jk}{}_{tr} + \delta^k_r H^{ij}{}_{st} + \delta^k_t H^{ij}{}_{rs}$$
$$+ \delta^k_s H^{ij}{}_{tr} + \delta^j_r H^{ki}{}_{st} + \delta^j_t H^{ki}{}_{rs} + \delta^j_s H^{ki}{}_{tr} = 0.$$

Hence show that, in $X_4$,

$$H^{ij}{}_{rs} H^{rs}{}_{kj} = \tfrac{1}{4}\delta^i_k (H^{lj}{}_{rs} H^{rs}{}_{lj}).$$

4.10  In $X_4$ and under the circumstances of Problem 4.9, show that if $A_i$, $B^j$ are such that

$$H^{ij}{}_{kl} A_i = 0, \qquad H^{ij}{}_{kl} B^k = 0$$

then either

$$H^{ij}{}_{kl} = 0 \quad \text{or} \quad A_i B^i = 0.$$

4.11  Show that in $X_4$ $a_{ij} = a_{ji}$ with $a = \det(a_{ij}) \neq 0$ and $b_{ij} = -b_{ji}$ then

$$\det|a_{ij} + b_{ij}| = a + b + \tfrac{1}{2} a^{hk} a^{ml} b_{hm} b_{kl}\, a,$$

where $b = \det(b_{ij})$ and $a^{hk} a_{kl} = \delta^h_l$.

4.12  Do Problem 3.37 using normal coordinates.

4.13  Establish the equivalence of (4.11) and (4.15) for a type (2, 3) relative tensor field of weight $-3$.

4.14  Show that

(a) $\pounds_v \delta^{i_1 \cdots i_r}_{j_1 \cdots j_r} = 0,$

(b) $\pounds_v \varepsilon^{i_1 \cdots i_n} = 0,$

(c) $\pounds_v v^i = 0.$

4.15  Prove that

$$\pounds_t \pounds_u v^i + \pounds_u \pounds_v t^i + \pounds_v \pounds_t u^i = 0.$$

4.16  If the connection is symmetric prove that

$$\pounds_v(u^i{}_{|j}) - (\pounds_v u^i)_{|j} = u^h \pounds_v \Gamma_j{}^i{}_h.$$

4.17  If the connection is symmetric prove that

$$(\pounds_v \Gamma_j{}^i{}_k)_{|h} - (\pounds_v \Gamma_j{}^i{}_h)_{|k} = \pounds_v K_j{}^i{}_{kh}.$$

4.18  Prove that

$$\pounds_u v^i = -\pounds_v u^i.$$

If $w^i = \pounds_u v^i$ show that

$$(\pounds_u \pounds_v - \pounds_v \pounds_u) T^{ij} = \pounds_w T^{ij}.$$

4.19  If $a_{ij} = a_{ji}$ and $\det(a_{ij}) \neq 0$ show that

$$\pounds_v a_{ij} = v_{i|j} + v_{j|i}$$

where $v_i = a_{ij} v^j$ and $\Gamma_i{}^h{}_j = \overset{(a)}{\Gamma}_i{}^h{}_j.$

4.20  Evaluate $\det(t_i{}^j)$ where $t_i{}^j = \lambda \delta^j_i + A^j B_i$.

4.21  Let $a^i_j$ be the components of an $n \times n$ matrix and define

$$M_{(0)h}{}^j = \delta^j_h, \qquad M_{(s)h}{}^j = a^l_h M_{(s-1)l}{}^j \quad \text{for} \quad s = 1, 2, \ldots.$$

Show that, if $a_{(s)}$ is defined by (2.34) then for $1 \leq r \leq n$

$$\delta^{j_1 j_2 \cdots j_r}_{h_1 h_2 \cdots h_r} a^{h_2}_{j_2} \cdots a^{h_r}_{j_r} = (r-1)! \sum_{s=0}^{r-1} (-1)^s M_{(s)h_1}{}^{j_1} a_{(r-s-1)},$$

and hence that

$$A^j_h = \sum_{s=0}^{n-1} (-1)^s M_{(s)h}{}^j a_{(n-s-1)},$$

where $A^j_h$ is the cofactor of $a^h_j$. Establish that

$$(-1)^n M_{(n)l}{}^j + \sum_{s=1}^{n} (-1)^{n-s} M_{(n-s)l}{}^j a_{(s)} = 0.$$

Show that this is just the Cayley-Hamilton theorem, by comparing it with (2.39), where $\lambda$ is replaced by $-\lambda$.

4.22  If, in $X_4$, the quantities $H^i_j$ are such that $H^i_j = \delta^i_j + \lambda A^i_j$ where $A^i_j H^j_k = \delta^i_k$, show that the condition $A^i_i = 0$ implies that $\det(H^i_j) = -27/16$.

# 5

# THE CALCULUS OF DIFFERENTIAL FORMS

The calculus of differential forms—often called exterior differential calculus—represents a relatively new tool of analysis whose use in pure and applied mathematics has become increasingly widespread.† Like the tensor calculus, its origins are to be found in differential geometry, largely as the result of the investigations of E. Cartan towards the beginning of this century. It was soon recognized, however, that the ramifications of this subject extend far beyond the realm of differential geometry, and accordingly it can properly be regarded as belonging to analysis per se: indeed, it has recently found its way into several modern texts on advanced differential and integral calculus. In these expositions, as well as in the books by H. Cartan [1] and H. Flanders [1], the exterior differential calculus is introduced and treated without reference to the tensor calculus, and it has even been asserted that the former is superior to the latter, this opinion being based on the fact that the use of forms is restricted to a lesser extent to coordinate neighborhoods than the use of tensors. It is, perhaps, somewhat more realistic to recognize that both disciplines are indispensable, each in its own right, there being many situations in which one is more effective than the other. In fact, some of the most significant successes have been achieved when both are used in conjunction with each other, a phenomenon which is exemplified most convincingly by the works of E. Cartan.

It is this second point of view which is reflected in this chapter. The

† Differential forms are also treated in the following texts: Brehmer and Haar [1], E. Cartan [3], H. Cartan [1], Flanders [1], and Slebodzinski [1]. For applications of differential forms to general relativity, see Israel [1]. Differential forms are also extremely useful in the theory of partial differential equations; see, for instance, Harrison and Estabrook [1] and Wahlquist and Estabrook [1].

algebraic and analytical properties of differential forms are treated in a somewhat intuitive manner based on the use of coordinates, the reader being referred to the Appendix for a more rigorous and intrinsic approach. Since the application of the exterior differential calculus is particularly effective whenever integrability conditions are involved, we have inserted a fairly detailed description of several versions of the theorem of Frobenius on systems of total differential equations, to which frequent appeal is made in the sequel.

One of the most powerful results of the exterior differential calculus is the so-called theorem of Stokes, which represents a generalization to manifolds of the Gauss divergence theorem of three-dimensional Euclidean geometry. It should be stressed, however, that in contrast to the theorem of Gauss, the theorem of Stokes is entirely independent of metric notions such as length, area, and volume. In a sense this theorem may be regarded as the fundamental theorem of calculus on manifolds, and consequently a rigorous derivation of its most general form depends on the theory of integration on manifolds which is beyond the scope of this volume. Nevertheless, the version of Stokes' theorem derived below should be sufficient to meet most requirements.

## 5.1 THE EXTERIOR (OR WEDGE) PRODUCT OF DIFFERENTIAL FORMS

Let $A_j$ denote the components of a covariant vector field on a differentiable manifold $X_n$ referred to local coordinates $x^j$. Corresponding to an arbitrary displacement on $X_n$ with components $dx^j$ we can construct the sum

$$\omega = A_j \, dx^j, \tag{1.1}$$

which is called a 1-*form* (or *Pfaffian* form). The 1-form (1.1) happens to be a scalar; however, given a type (1, 1) tensor field $A_j^h$ on $X_n$, one can similarly define other 1-forms by constructing the expressions $A_j^h \, dx^j$, which we shall regard as contravariant 1-forms. In particular, with $A_j^h = \delta_j^h$, we obtain the 1-forms $dx^h$.

We shall begin with a brief consideration of purely algebraic operations which may be performed on 1-forms. For instance, the sum of two scalar 1-forms $\omega = A_j \, dx^j$, $\pi = B_j \, dx^j$ is defined by the obvious rule

$$\omega + \pi = (A_j + B_j) \, dx^j,$$

which is again a 1-form; needless to say, the sum of two 1-forms of different types, such as a scalar and a contravariant 1-form, has no tensorial meaning. Similarly, the product of a 1-form $\omega$ as exemplified by (1.1) and any function $f(x^h)$ which depends only on the coordinates is defined to be

$$f\omega = fA_j \, dx^j, \tag{1.2}$$

which is another 1-form. However, when the product of two 1-forms $\omega$, $\pi$ is considered, a new concept must be introduced, namely, the so-called *exterior* (or *wedge*) product of $\omega$ and $\pi$, to be denoted by $\omega \wedge \pi$. The exterior product is assumed to obey the usual distributive law of multiplication, but instead of commutativity one stipulates anticommutativity for 1-forms:

$$\omega \wedge \pi = -\pi \wedge \omega. \tag{1.3}$$

Thus, in particular, we have

$$dx^j \wedge dx^h = -dx^h \wedge dx^j. \tag{1.4}$$

As a result of (1.2) and these rules, the exterior product of two scalar 1-forms $\omega = A_j\, dx^j$, $\pi = B_h\, dx^h$, can be expressed in the form

$$\omega \wedge \pi = (A_j\, dx^j) \wedge (B_h\, dx^h) = A_j B_h\, dx^j \wedge dx^h, \tag{1.5}$$

or

$$\omega \wedge \pi = \tfrac{1}{2}(A_j B_h\, dx^j \wedge dx^h + A_h B_j\, dx^h \wedge dx^j) = \tfrac{1}{2}(A_j B_h - A_h B_j)\, dx^j \wedge dx^h. \tag{1.6}$$

This is not a 1-form but represents a scalar 2-form, the most general 2-form of this type being represented by expressions such as

$$A_{jh}\, dx^j \wedge dx^h, \tag{1.7}$$

where, because of (1.4), it can be assumed without loss of generality that the coefficients $A_{jh}$ are the components of a skew-symmetric type (0, 2) tensor.

It will be seen that the process of exterior multiplication is extremely useful; one of its main motivations arose from the transformation law of integrals under a change of independent variables. For instance, for a coordinate transformation $\bar{x}^j = \bar{x}^j(x^h)$ with positive Jacobian we have

$$d\bar{x}^j = \frac{\partial \bar{x}^j}{\partial x^h}\, dx^h, \tag{1.8}$$

so that

$$d\bar{x}^j \wedge d\bar{x}^l = \frac{\partial \bar{x}^j}{\partial x^h}\frac{\partial \bar{x}^l}{\partial x^k}\, dx^h \wedge dx^k = \frac{1}{2}\frac{\partial(\bar{x}^j, \bar{x}^l)}{\partial(x^h, x^k)}\, dx^h \wedge dx^k. \tag{1.9}$$

Thus, when $n = 2$, the left-hand side consists of only one nonvanishing term, and the two nonzero terms on the right can be combined to yield

$$d\bar{x}^1 \wedge d\bar{x}^2 = \frac{\partial(\bar{x}^1, \bar{x}^2)}{\partial(x^1, x^2)}\, dx^1 \wedge dx^2.$$

In a purely formal sense this leads directly to the correct formula for the transformation of the integral of a scalar function $g(\bar{x}^j)$ over a domain $G$

of our 2-dimensional manifold $X_2$ provided that we represent the integrand as a 2-form:

$$\int_G g(\bar{x}^j)\, d\bar{x}^1 \wedge d\bar{x}^2 = \int_G g(\bar{x}^j(x^h))\frac{\partial(\bar{x}^1, \bar{x}^2)}{\partial(x^1, x^2)}\, dx^1 \wedge dx^2, \tag{1.10}$$

it being recalled that the Jacobian on the right-hand side is supposed to be positive.

The exterior product of more than two 1-forms can obviously be obtained recursively, it being assumed that the associative law holds: for $p$ 1-forms, with $p \leq n$, given by

$$\omega^r = A_j^r\, dx^j \qquad (r = 1, \ldots, p \leq n), \tag{1.11}$$

we have

$$\left.\begin{aligned} \omega^1 \wedge \omega^2 \wedge \omega^3 &= (\omega^1 \wedge \omega^2) \wedge \omega^3 = A_{j_1}^1 A_{j_2}^2 A_{j_3}^3\, dx^{j_1} \wedge dx^{j_2} \wedge dx^{j_3}, \\ \hdotsfor{1} \\ \omega^1 \wedge \cdots \wedge \omega^p &= A_{j_1}^1 \cdots A_{j_p}^p\, dx^{j_1} \wedge \cdots \wedge dx^{j_p}. \end{aligned}\right\} \tag{1.12}$$

As in the case of (1.6) these exterior products can be written in slightly more illuminating form. From the property (4.2.15) of the generalized Kronecker delta it follows immediately that for any set of quantities $C^{j_1 \cdots j_p}$ which are completely skew-symmetric in their superscripts, one has

$$\delta_{h_1 \cdots h_p}^{j_1 \cdots j_p} C^{h_1 \cdots h_p} = p!\, C^{j_1 \cdots j_p}. \tag{1.13}$$

In particular,

$$dx^{j_1} \wedge \cdots \wedge dx^{j_p} = \frac{1}{p!} \delta_{h_1 \cdots h_p}^{j_1 \cdots j_p}\, dx^{h_1} \wedge \cdots \wedge dx^{h_p}, \tag{1.14}$$

so that (1.12) is equivalent to

$$\left.\begin{aligned} \omega^1 \wedge \omega^2 \wedge \omega^3 &= \frac{1}{3!} \delta_{h_1 h_2 h_3}^{j_1 j_2 j_3} A_{j_1}^1 A_{j_2}^2 A_{j_3}^3\, dx^{h_1} \wedge dx^{h_2} \wedge dx^{h_3}, \\ \hdotsfor{1} \\ \omega^1 \wedge \cdots \wedge \omega^p &= \frac{1}{p!} \delta_{h_1 \cdots h_p}^{j_1 \cdots j_p} A_{j_1}^1 \cdots A_{j_p}^p\, dx^{h_1} \wedge \cdots \wedge dx^{h_p}. \end{aligned}\right\} \tag{1.15}$$

Here it is to be noted that $\delta_{h_1 \cdots h_p}^{j_1 \cdots j_p} A_{j_1}^1 \cdots A_{j_p}^p$ is the determinant of the matrix obtained by choosing the $h_1$th, ..., $h_p$th columns of the $p \times n$ matrix

$$\begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & \cdots & A_n^2 \\ \hdotsfor{4} \\ A_1^p & A_2^p & \cdots & A_n^p \end{pmatrix}. \tag{1.16}$$

In particular, for $p = n$, the determinant $\pm \det(A_j^h)$ appears $n!$ times in $\omega^1 \wedge \cdots \wedge \omega^n$, so that

$$\omega^1 \wedge \cdots \wedge \omega^n = \det(A_j^h) \, dx^1 \wedge \cdots \wedge dx^n. \tag{1.17}$$

The $p$ 1-forms (1.11) are said to be linearly independent if the $p$ quantities whose components are $A_j^1, \ldots, A_j^p$ are linearly independent. From the above construction it is immediately evident that *the 1-forms $\omega^1, \ldots, \omega^p$ are linearly dependent if and only if*

$$\omega^1 \wedge \cdots \wedge \omega^p = 0. \tag{1.18}$$

Clearly there cannot exist more than $n$ linearly independent 1-forms at any point $P$ of $X_n$. Thus the scalar 1-forms at $P$ constitute a vector space of dimension $n$, whose basis elements are represented by $dx^1, \ldots, dx^n$ relative to the given coordinate system of $X_n$. Since each scalar 1-form is uniquely defined by a covariant vector $A_j$ as in (1.1), this vector space is isomorphic to the dual tangent space $T_n^*(P)$ of covariant vectors of $X_n$ at $P$.

Similarly all scalar 2-forms (1.7) at $P$ are the elements of a vector space of dimension $\frac{1}{2}n(n-1)$, whose basis elements are represented by the $\frac{1}{2}n(n-1)$ exterior products $dx^j \wedge dx^h$ with $j < h$. [It is easily verified that these basis elements are linearly independent. For, suppose to the contrary that there exist parameters $c_{jh}$, not all zero, such that

$$\sum_{j < h} c_{jh} \, dx^j \wedge dx^h = 0.$$

The exterior product of this relation with $dx^3 \wedge \cdots \wedge dx^n$ is formed, and, since $dx^r \wedge dx^r = 0$ by virtue of (1.4), it is clear that the sole surviving term is

$$c_{12} \, dx^1 \wedge dx^2 \wedge dx^3 \wedge \cdots \wedge dx^n,$$

from which it follows that $c_{12} = 0$. By symmetry, we conclude that all $c_{jh}$ vanish, which establishes a contradiction.] Continuing in this manner, we regard any scalar $p$-form defined by

$$\omega = A_{j_1 \ldots j_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \tag{1.19}$$

where $A_{j_1 \ldots j_p}$ is a completely skew-symmetric type $(0, p)$ tensor with $p \leq n$, as an element of a vector space of dimension $\binom{n}{p}$, whose basis elements relative to the given coordinate system on $X_n$ are represented by $dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_p}$, with $j_1 < j_2 < \cdots < j_p$. (The direct sum of these vector spaces ($p = 0, 1, \ldots, n$) is said to constitute a *Grassmann algebra*.)

Although it is not possible to form the sum of a $p$-form $\omega$ and a $q$-form $\pi$ unless $p = q$, one can always construct their exterior product. For a $p$-form $\omega$ given by (1.19), and $q$-form $\pi$ given by

$$\pi = B_{h_1 \ldots h_q} \, dx^{h_1} \wedge \cdots \wedge dx^{h_q}, \tag{1.20}$$

the exterior product is defined by

$$\omega \wedge \pi = A_{j_1 \ldots j_p} B_{h_1 \ldots h_q} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^{h_1} \wedge \cdots \wedge dx^{h_q}. \tag{1.21}$$

In connection with this definition it should be observed that

$$dx^{j_1} \wedge \cdots \wedge dx^{j_{p-1}} \wedge dx^{j_p} \wedge dx^{h_1} \wedge \cdots \wedge dx^{h_q}$$
$$= (-1)^q \, dx^{j_1} \wedge \cdots \wedge dx^{j_{p-1}} \wedge dx^{h_1} \wedge \cdots \wedge dx^{h_q} \wedge dx^{j_p},$$

since the rule (1.4) is applied $q$ times in succession in the process of moving $dx^{j_p}$ from the $p$th to the $(p + q)$th position. Carrying out this operation $p$ times, we obtain

$$dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^{h_1} \wedge \cdots \wedge dx^{h_q}$$
$$= (-1)^{pq} \, dx^{h_1} \wedge \cdots \wedge dx^{h_q} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p}. \tag{1.22}$$

When this result is applied to the definition (1.21), we obtain the following rule, namely,

$$\omega \wedge \pi = (-1)^{pq} \pi \wedge \omega, \tag{1.23}$$

for the exterior product of a $p$-form $\omega$ with a $q$-form $\pi$. Clearly (1.3), in which $p = q = 1$, is a special case of this.

For future reference we note the following. Suppose that, for a set of quantities $A_{jh}$ endowed with two subscripts (not necessarily skew-symmetric), we are given that the corresponding 2-form vanishes:

$$\omega = A_{jh} \, dx^j \wedge dx^h = 0. \tag{1.24}$$

Clearly one cannot infer from this that $A_{jh} = 0$, for the symmetric part (if any) of $A_{jh}$ is excluded from this statement because of (1.4). However, we can write

$$\omega = \sum_{j < h} (A_{jh} - A_{hj}) \, dx^j \wedge dx^h, \tag{1.25}$$

and hence, because the basis elements $dx^j \wedge dx^h$ with $j < h$ are linearly independent, we may infer that

$$A_{jh} - A_{hj} = \delta_{jh}^{lk} A_{lk} = 0. \tag{1.26}$$

More generally, let us consider a set of quantities $A_{h_1 \ldots h_p}$ endowed with $p$ subscripts (not necessarily skew-symmetric). From (1.14) we have that

$$A_{h_1 \ldots h_p} \, dx^{h_1} \wedge \cdots \wedge dx^{h_p} = \frac{1}{p!} \delta_{j_1 \ldots j_p}^{h_1 \ldots h_p} A_{h_1 \ldots h_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}. \tag{1.27}$$

Moreover,

$$\sum_{j_1 < \cdots < j_p} \delta_{j_1 \ldots j_p}^{h_1 \ldots h_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p} = \frac{1}{p!} \delta_{j_1 \ldots j_p}^{h_1 \ldots h_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \tag{1.28}$$

for, among the $p!$ permutations of $p$ distinct integers $j_1, \ldots, j_p$, there is only one for which $j_1 < j_2 < \cdots < j_p$. Thus (1.27) can be written in the form

$$A_{h_1 \cdots h_p} \, dx^{h_1} \wedge \cdots \wedge dx^{h_p} = \sum_{j_1 < \cdots < j_p} \delta_{j_1 \cdots j_p}^{h_1 \cdots h_p} A_{h_1 \cdots h_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \quad (1.29)$$

which is valid for any $A_{h_1 \cdots h_p}$, irrespective of symmetry or skew-symmetry properties. In particular, it now follows from the linear independence of the basis elements $dx^{j_1} \wedge \cdots \wedge dx^{j_p}$ with $j_1 < j_2 < \cdots < j_p$, that the condition

$$A_{h_1 \cdots h_p} \, dx^{h_1} \wedge \cdots \wedge dx^{h_p} = 0 \quad (1.30)$$

implies that

$$\delta_{j_1 \cdots j_p}^{h_1 \cdots h_p} A_{h_1 \cdots h_p} = 0. \quad (1.31)$$

This fact will be used frequently in the sequel.

## 5.2  EXTERIOR DERIVATIVES OF $p$-FORMS

Having briefly dealt with the algebraic properties of $p$-forms, we shall now turn to the differential calculus of such forms. This is based on the concept of *exterior derivative*, which is to be described below.

For a 1-form $\omega$ given by (1.1) the exterior derivative, denoted by $d\omega$, is a 2-form defined by

$$d\omega = \frac{\partial A_j}{\partial x^k} \, dx^k \wedge dx^j. \quad (2.1)$$

Because of (1.4) this can be written in the form

$$d\omega = -\frac{\partial A_j}{\partial x^k} \, dx^j \wedge dx^k. \quad (2.2)$$

Incidentally, one can combine (2.1) and (2.2) to obtain

$$d\omega = \frac{1}{2}\left(\frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k}\right) dx^j \wedge dx^k, \quad (2.3)$$

which shows that the coefficients of the basis elements $dx^j \wedge dx^k$ with $j < k$ are the components of the "curl" of the vector field $A_j$.

Similarly, the exterior derivative of the 2-form (1.7) is, by definition, the 3-form

$$d\omega = \frac{\partial A_{jh}}{\partial x^k} \, dx^k \wedge dx^j \wedge dx^h = \frac{\partial A_{jh}}{\partial x^k} \, dx^j \wedge dx^h \wedge dx^k. \quad (2.4)$$

In general, then, for a $p$-form $(1 \le p \le n)$

$$\omega = A_{j_1 \cdots j_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \quad (2.5)$$

the *exterior derivative* $d\omega$ is the $(p + 1)$-form

$$d\omega = \frac{\partial A_{j_1 \cdots j_p}}{\partial x^k} \, dx^k \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p}$$

$$= (-1)^p \frac{\partial A_{j_1 \cdots j_p}}{\partial x^{j_{p+1}}} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^{j_{p+1}}. \quad (2.6)$$

We shall now briefly derive some basic properties of the process of exterior differentiation. Clearly, the exterior derivative of the sum of two $p$-forms is simply the sum of their exterior derivatives. However, for the exterior product of the $p$-form (1.19) with the $q$-form (1.20), we find, on applying the definition (2.6) to the product (1.21):

$$d(\omega \wedge \pi) = \left[\frac{\partial A_{j_1 \cdots j_p}}{\partial x^k} \, dx^k B_{h_1 \cdots h_q} + A_{j_1 \cdots j_p} \frac{\partial B_{h_1 \cdots h_q}}{\partial x^k} \, dx^k\right]$$
$$\wedge \, dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^{h_1} \wedge \cdots \wedge dx^{h_q},$$

or, because

$$dx^k \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p} = (-1)^p \, dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^k,$$

this expression can be written in the form

$$d(\omega \wedge \pi) = \left(\frac{\partial A_{j_1 \cdots j_p}}{\partial x^k} \, dx^k \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p}\right) \wedge (B_{h_1 \cdots h_q} \, dx^{h_1} \wedge \cdots \wedge dx^{h_q})$$

$$+ (-1)^p (A_{j_1 \cdots j_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}) \wedge \left(\frac{\partial B_{h_1 \cdots h_q}}{\partial x^k} \, dx^k \wedge dx^{h_1} \wedge \cdots \wedge dx^{h_q}\right).$$

From (1.19), (1.20), and the definition (2.6) it therefore follows that

$$d(\omega \wedge \pi) = d\omega \wedge \pi + (-1)^p \omega \wedge d\pi, \quad (2.7)$$

which represents the *product rule for the exterior derivative of the exterior product of a p-form $\omega$ with a q-form $\pi$*. It is easily verified that an interchange of the order of $\omega$ and $\pi$ in (2.7) is consistent with (1.23).

Another immediate consequence of the definition (2.6) is the fact that

$$d(d\omega) = 0 \quad (2.8)$$

*identically for any p-form (2.5) provided that the coefficients $A_{j_1 \cdots j_p}$ in (2.5) are of class $C^2$.* For, if the definition (2.6) is applied to itself it is found that

$$d(d\omega) = \frac{\partial^2 A_{j_1 \cdots j_p}}{\partial x^h \, \partial x^k} \, dx^h \wedge dx^k \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p},$$

which vanishes identically by virtue of the symmetry of the second derivatives of $A_{j_1 \cdots j_p}$ in $h$ and $k$. The identity (2.8) is often referred to as *Poincaré's Lemma*; we shall see subsequently that this lemma possesses a very useful local inverse.

Let us suppose, for the moment, that the coefficients $A_{j_1 \cdots j_p}$ in (2.5) are the components of a type $(0, p)$ tensor, so that the $p$-form (2.5) is a scalar. However, it was repeatedly stressed that the partial derivatives with respect to $x^k$ of a tensor are not, in general, tensorial. Hence it is not immediately evident that the exterior derivative (2.6) is tensorial. In order to clarify this point, we apply the identity (1.27) to the $(p + 1)$-form (2.6), which yields the following expression:

$$d\omega = \frac{(-1)^p}{(p+1)!} \delta_{j_1 \cdots j_p j_{p+1}}^{h_1 \cdots h_p h_{p+1}} \frac{\partial A_{h_1 \cdots h_p}}{\partial x^{h_{p+1}}} dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^{j_{p+1}}. \quad (2.9)$$

We now recall the general result obtained in Section 4.2, according to which the sum

$$\delta_{j_1 \cdots j_p j_{p+1}}^{h_1 \cdots h_p h_{p+1}} \frac{\partial A_{h_1 \cdots h_p}}{\partial x^{h_{p+1}}} \quad (2.10)$$

is a type $(0, p + 1)$ tensor whenever $A_{h_1 \cdots h_p}$ represents a type $(0, p)$ tensor field; it therefore follows that, under these circumstances, *the exterior derivative* (2.9) *is a scalar* $(p + 1)$-*form*. At this juncture it should be stressed that the tensorial character of (2.10) is established without reference to the existence of a connection on our underlying differentiable manifold $X_n$.

However, in practice one is frequently confronted with $p$-forms which are not scalars. For instance, let us consider a contravariant 2-form of the type

$$\Pi^j = A_h{}^j{}_k \, dx^h \wedge dx^k, \quad (2.11)$$

in which the coefficients $A_h{}^j{}_k$ represent the components of a type $(1, 2)$ tensor field. According to (2.6) and (2.9) we have

$$d\Pi^j = \frac{\partial A_h{}^j{}_k}{\partial x^l} dx^l \wedge dx^h \wedge dx^k = \frac{1}{3!} \delta_{hkl}^{rst} \frac{\partial A_r{}^j{}_s}{\partial x^t} dx^h \wedge dx^k \wedge dx^l, \quad (2.12)$$

which is not in general tensorial. In order to construct a tensorial exterior derivative we have to invoke a connection, and accordingly it is now assumed that $X_n$ is endowed with connection coefficients denoted by $\Gamma_h{}^j{}_k$. This allows us to introduce the covariant derivative of $A_h{}^j{}_k$, which, according to Section 3.5, is given by

$$A_r{}^j{}_{s|t} = \frac{\partial A_r{}^j{}_s}{\partial x^t} + A_r{}^m{}_s \Gamma_m{}^j{}_t - A_m{}^j{}_s \Gamma_r{}^m{}_t - A_r{}^j{}_m \Gamma_s{}^m{}_t. \quad (2.13)$$

Because of the skew-symmetry properties of the generalized Kronecker delta we have

$$\delta_{hkl}^{rst}(A_m{}^j{}_s \Gamma_r{}^m{}_t + A_r{}^j{}_m \Gamma_s{}^m{}_t) = \tfrac{1}{2} \delta_{hkl}^{rst}(A_m{}^j{}_s S_r{}^m{}_t + A_r{}^j{}_m S_s{}^m{}_t), \quad (2.14)$$

where $S_r{}^m{}_t$ denotes the torsion tensor of our connection:

$$S_r{}^m{}_t = \Gamma_r{}^m{}_t - \Gamma_t{}^m{}_r. \quad (2.15)$$

Thus, when (2.14) is applied to (2.13), we obtain

$$\delta_{hkl}^{rst}[A_r{}^j{}_{s|t} + \tfrac{1}{2}(A_m{}^j{}_s S_r{}^m{}_t + A_r{}^j{}_m S_s{}^m{}_t)] = \delta_{hkl}^{rst}\left(\frac{\partial A_r{}^j{}_s}{\partial x^t} + A_r{}^m{}_s \Gamma_m{}^j{}_t\right), \quad (2.16)$$

from which it is evident that the right-hand side is a type $(1, 3)$ tensor. This conclusion, taken in conjunction with (2.12), strongly suggests that we should define the *covariant exterior derivative* of the contravariant 2-form (2.11) by

$$D\Pi^j = \frac{1}{3!} \delta_{hkl}^{rst}\left(\frac{\partial A_r{}^j{}_s}{\partial x^t} + A_r{}^m{}_s \Gamma_m{}^j{}_t\right) dx^h \wedge dx^k \wedge dx^l$$

$$= d\Pi^j + \frac{1}{3!} \delta_{hkl}^{rst} A_r{}^m{}_s \Gamma_m{}^j{}_t \, dx^h \wedge dx^k \wedge dx^l, \quad (2.17)$$

which is a contravariant 3-form. Because of (2.16) this relation is equivalent to

$$D\Pi^j = \frac{1}{3!} \delta_{hkl}^{rst}[A_r{}^j{}_{s|t} + \tfrac{1}{2}(A_m{}^j{}_s S_r{}^m{}_t + A_r{}^j{}_m S_s{}^m{}_t)] dx^h \wedge dx^k \wedge dx^l. \quad (2.18)$$

The definition (2.17) can be written in a somewhat more illuminating form as follows. On applying the general identity (1.27) to the second term on the right-hand side of (2.17), after which we substitute from (2.11), we obtain

$$\frac{1}{3!} \delta_{hkl}^{rst} A_r{}^m{}_s \Gamma_m{}^j{}_t \, dx^h \wedge dx^k \wedge dx^l$$

$$= A_h{}^m{}_k \Gamma_m{}^j{}_l \, dx^h \wedge dx^k \wedge dx^l = \Pi^h \wedge \Gamma_h{}^j{}_l \, dx^l. \quad (2.19)$$

We therefore define the following 1-form (in the notation of E. Cartan):

$$\omega_h{}^j = \Gamma_h{}^j{}_l \, dx^l, \quad (2.20)$$

where it should be stressed that this form is *not* tensorial, being defined directly by means of the coefficients of our connection. By means of (2.19) and (2.20) we can now express the *covariant exterior derivative* (2.17) *of the* 2-*form* (2.11) *in the following final form*:

$$D\Pi^j = d\Pi^j + \Pi^h \wedge \omega_h{}^j = d\Pi^j + \omega_h{}^j \wedge \Pi^h, \quad (2.21)$$

where we have used (1.23).

By means of a similar application of (1.27) to (2.18) it is found that (2.21) is equivalent to

$$D\Pi^j = [A_{h}{}^{j}{}_{k|l} + \tfrac{1}{2}(A_{m}{}^{j}{}_{k}S_{h}{}^{m}{}_{l} + A_{h}{}^{j}{}_{m}S_{k}{}^{m}{}_{l})] \, dx^h \wedge dx^k \wedge dx^l. \qquad (2.22)$$

In passing we note that, should $A_{h}{}^{j}{}_{m} = -A_{m}{}^{j}{}_{h}$, this expression reduces to

$$D\Pi^j = (A_{h}{}^{j}{}_{k|l} + A_{m}{}^{j}{}_{k}S_{h}{}^{m}{}_{l}) \, dx^h \wedge dx^k \wedge dx^l. \qquad (2.23)$$

The same technique may be applied to the more general $p$-form defined by

$$\Pi^j = A^{j}_{j_1 \cdots j_p} \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \qquad (2.24)$$

in which $A^{j}_{j_1 \cdots j_p}$ denotes the components of a type $(1, p)$ tensor. The *covariant exterior derivative of this $p$-form is given by the following contravariant $(p + 1)$-form*:

$$D\Pi^j = d\Pi^j + (-1)^p \Pi^h \wedge \omega_{h}{}^{j} = d\Pi^j + \omega_{h}{}^{j} \wedge \Pi^h, \qquad (2.25)$$

of which (2.21) is a special case. The extension of this formula to $p$-forms whose coefficients are type $(r, p)$ tensors is immediate. However, one is sometimes confronted with $p$-forms whose coefficients are type $(r, p + q)$ tensors $(q \geq 1)$, in which case an additional $q$ terms which are linear in the 1-forms (2.20) enter into the definition of the covariant exterior derivative. For instance, in the case of a 2-form

$$\Pi_{l}{}^{j} = A_{l}{}^{j}{}_{hk} \, dx^h \wedge dx^k, \qquad (2.26)$$

whose coefficients are the components of a type $(1, 3)$ tensor field, the exterior covariant derivative must be defined as

$$D\Pi_{l}{}^{j} = d\Pi_{l}{}^{j} + \omega_{h}{}^{j} \wedge \Pi_{l}{}^{h} - \omega_{l}{}^{h} \wedge \Pi_{h}{}^{j}. \qquad (2.27)$$

It is easily verified that this may also be expressed in the form

$$D\Pi_{l}{}^{j} = [A_{l}{}^{j}{}_{hk|p} + \tfrac{1}{2}(A_{l}{}^{j}{}_{mk}S_{h}{}^{m}{}_{p} + A_{l}{}^{j}{}_{hm}S_{k}{}^{m}{}_{p})] \, dx^h \wedge dx^k \wedge dx^p, \qquad (2.28)$$

or, if $A_{l}{}^{j}{}_{hk}$ is skew-symmetric in $h$ and $k$,

$$D\Pi_{l}{}^{j} = (A_{l}{}^{j}{}_{hk|p} + A_{l}{}^{j}{}_{mk}S_{h}{}^{m}{}_{p}) \, dx^h \wedge dx^k \wedge dx^p. \qquad (2.29)$$

In particular, if one regards the components $X^j$ of a type $(1, 0)$ tensor field as representing $n$ 0-forms, one may write

$$DX^j = dX^j + \omega_{h}{}^{j}X^h, \qquad (2.30)$$

which, by virtue of (2.20), coincides with the absolute differential of $X^j$ (indicating that the notation introduced here is consistent with that of Section 3.4).

In conclusion we observe that the product rule for covariant exterior differentiation is the same as the rule (2.7) for exterior differentiation: if

$\Omega$, $\Pi$ respectively denote a $p$-form and a $q$-form of arbitrary types (the tensor indices being suppressed), we have

$$D(\Omega \wedge \Pi) = D\Omega \wedge \Pi + (-1)^p \Omega \wedge D\Pi. \qquad (2.31)$$

The proof of this assertion follows directly from (2.7) and the general definition of exterior covariant derivatives.

## 5.3   THE LEMMA OF POINCARÉ AND ITS CONVERSE

Let $\omega$ be a $p$-form given by (2.5). This form is said to be *exact*, if there exists a $(p - 1)$-form $\pi$ such that

$$\omega = d\pi. \qquad (3.1)$$

According to the Lemma of Poincaré [i.e., (2.8)],

$$d\omega = 0 \qquad (3.2)$$

whenever $\omega$ is exact. Any $p$-form $\omega$ satisfying the condition (3.2) is called *closed*, and thus the Lemma of Poincaré states that *every exact $p$-form is closed*. This simple statement contains as special cases some well-known identities of elementary three-dimensional vector analysis.

For instance, let us suppose that $\omega$ is the 1-form (1.1). By means of (2.3) we can then write

$$d\omega = \tfrac{1}{2}F_{jk} \, dx^j \wedge dx^k, \qquad (3.3)$$

where we have put

$$F_{jk} = \frac{\partial A_k}{\partial x^j} - \frac{\partial A_j}{\partial x^k}. \qquad (3.4)$$

Now, let us suppose for the moment that (1.1) is exact, that is, that there exists a scalar function $\phi(x^h) \in C^2$ such that

$$\omega = A_j \, dx^j = d\phi = \frac{\partial \phi}{\partial x^j} \, dx^j, \qquad (3.5)$$

which implies that

$$A_j = \frac{\partial \phi}{\partial x^j}. \qquad (3.6)$$

Under these circumstances the covariant vector field $A_j$ is the gradient field of $\phi(x^h)$, while $d(d\phi) = d\omega$ as given by (3.3) vanishes. Thus $F_{jk} = 0$ [as would also be evident directly by substitution of (3.6) in (3.4)], which is tantamount to the well-known statement that curl(grad $\phi$) = 0 identically (irrespective of dimension).

Moreover, even when $F_{jk} \neq 0$, an application of the Lemma of Poincaré to (3.3) gives

$$0 = d(d\omega) = \frac{1}{2} \frac{\partial F_{jk}}{\partial x^h} dx^h \wedge dx^j \wedge dx^k,$$

and according to the final remark of Section 5.1 it is inferred that

$$\delta_{jkh}^{rst} \frac{\partial F_{rs}}{\partial x^t} = 0,$$

or

$$\frac{\partial F_{jk}}{\partial x^h} + \frac{\partial F_{hj}}{\partial x^k} + \frac{\partial F_{kh}}{\partial x^j} = 0, \qquad (3.7)$$

[which can also be verified directly by differentiation of the definition (3.4) of $F_{jk}$]. Now, for $n = 3$, this is merely a single relation, namely

$$\frac{\partial F_{23}}{\partial x^1} + \frac{\partial F_{31}}{\partial x^2} + \frac{\partial F_{12}}{\partial x^3} = 0, \qquad (3.8)$$

in which, according to the usual definition, $F_{23}, F_{31}$, and $F_{12}$ are respectively the $x^1$-, $x^2$-, $x^3$-components of the curl of the vector $\mathbf{A}$ whose components are given by $A_1, A_2, A_3$. Thus (3.7) is equivalent to the well-known identity div(curl $\mathbf{A}$) = 0 of three-dimensional vector analysis.

The main object of the present section is an investigation into the validity or otherwise of the converse of the Lemma of Poincaré, which would be contained in a statement to the effect that every closed $p$-form is exact, or equivalently, that the relation $d\omega = 0$ for a given $p$-form $\omega$ implies the existence of a $(p - 1)$-form $\pi$ such that $\omega = d\pi$. We shall see that this is indeed the case in a certain local sense to be specified below.

To this end, let us consider a $p$-form

$$\omega = A_{j_1 \cdots j_p} dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \qquad (3.9)$$

of which it is assumed that the coefficients $A_{j_1 \cdots j_p} = A_{j_1 \cdots j_p}(x^h)$ are class $C^1$ functions of the local coordinates $x^h$ of our differentiable manifold $X_n$. More precisely, for a given point $P$ on $X_n$ let us choose our coordinates such that $x^1 = \cdots = x^n = 0$ at $P$, after which we construct an open set $U$ on $X_n$ which is defined by the property that for any point $Q \in U$ with coordinates $x^h$, the segment consisting of the points with coordinates $tx^h$, $0 \leq t \leq 1$, is also contained in $U$. The set $U$ is said to be *star-shaped* with respect to the point $P$. It will henceforth be assumed that the coefficients $A_{j_1 \cdots j_p}$ in (3.9) are defined and of class $C^1$ on such a star-shaped set. Under these conditions

we can apply an operator $\mathcal{O}$ to our $p$-form $\omega$, this operator being defined by

$$\mathcal{O}\omega = \sum_{r=1}^{p} (-1)^{r-1} \left\{ \int_0^1 t^{p-1} A_{j_1 \cdots j_p}(tx^h) \, dt \right\} x^{j_r} dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}}$$
$$\wedge \ dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_p}. \qquad (3.10)$$

This is a $(p - 1)$-form, whose existence is assured by virtue of our assumptions. It has the general property that if $\omega = 0$ then $\mathcal{O}\omega = 0$, which is easily seen by expressing (3.10) in the form

$$\mathcal{O}\omega = \frac{1}{(p-1)!} \left\{ \int_0^1 t^{p-1} A_{j_1 \cdots j_p}(tx^h) \, dt \right\} x^{h_1} \delta_{h_1 \cdots h_p}^{j_1 \cdots j_p} dx^{h_2} \wedge \cdots \wedge dx^{h_p}$$

and noting (1.30) and (1.31).

We now evaluate the exterior derivative of $\mathcal{O}\omega$. In doing so, we observe that

$$\frac{\partial}{\partial x^j} A_{j_1 \cdots j_p}(tx^h) = \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^l} \cdot \frac{\partial(tx^l)}{\partial x^j} = \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^l} t\delta_j^l = t \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^j},$$

so that exterior differentiation of (3.10) gives

$$d(\mathcal{O}\omega) = \sum_{r=1}^{p} (-1)^{r-1} \left\{ \int_0^1 t^p \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^j} \, dt \right\} x^{j_r} dx^j$$
$$\wedge \ dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}} \wedge dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_p}$$
$$+ p \left\{ \int_0^1 t^{p-1} A_{j_1 \cdots j_p}(tx^h) \, dt \right\} dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \qquad (3.11)$$

where, in the last term on the right-hand side, it has been noted that

$$\sum_{r=1}^{p} (-1)^{r-1} dx^{j_r} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}}$$
$$\wedge \ dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_p} = p \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}.$$

Now let us apply the $\mathcal{O}$-operator to some $(p + 1)$-form $\Omega$, the latter being given by

$$\Omega = B_{j_1 \cdots j_{p+1}} dx^{j_1} \wedge \cdots \wedge dx^{j_{p+1}}, \qquad (3.12)$$

it being assumed once more that the coefficients $B_{j_1 \cdots j_{p+1}}$ are defined and of class $C^1$ on the star-shaped open set $U$. According to the definition (3.10)

of $\mathcal{O}$ we have

$$\mathcal{O}\Omega = \sum_{r=1}^{p+1} (-1)^{r-1} \left\{ \int_0^1 t^p B_{j_1 \cdots j_{p+1}}(tx^h) \, dt \right\} x^{j_r} dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}}$$
$$\wedge dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_{p+1}}$$

$$= \sum_{r=1}^{p} (-1)^{r-1} \left\{ \int_0^1 t^p B_{j_1 \cdots j_{p+1}}(tx^h) \, dt \right\} x^{j_r} dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}}$$
$$\wedge dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_{p+1}}$$

$$+ (-1)^p \left\{ \int_0^1 t^p B_{j_1 \cdots j_{p+1}}(tx^h) \, dt \right\} x^{j_{p+1}} dx^{j_1} \wedge \cdots \wedge dx^{j_p}.$$

In both terms on the right-hand side we replace the indices $j_{p+1}$ by $j$ and in the first term we move $dx^{j_{p+1}} = dx^j$ from the $p$th position to the first. Since this requires $(p-1)$ interchanges, we thus obtain

$$\mathcal{O}\Omega = (-1)^{p-1} \sum_{r=1}^{p} (-1)^{r-1} \left\{ \int_0^1 t^p B_{j_1 \cdots j_p j}(tx^h) \, dt \right\} x^{j_r} dx^j$$
$$dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}} \wedge dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_p}$$

$$+ (-1)^p \left\{ \int_0^1 t^p B_{j_1 \cdots j_p j}(tx^h) \, dt \right\} x^j dx^{j_1} \wedge \cdots \wedge dx^{j_p}. \qquad (3.13)$$

However, from (3.9) we have

$$d\omega = (-1)^p \frac{\partial A_{j_1 \cdots j_p}}{\partial x^{j_{p+1}}} dx^{j_1} \wedge \cdots \wedge dx^{j_p} \wedge dx^{j_{p+1}},$$

and thus we can make the identification

$$\Omega = d\omega, \qquad (3.14)$$

provided that we now put

$$B_{j_1 \cdots j_p j_{p+1}} = (-1)^p \frac{\partial A_{j_1 \cdots j_p}}{\partial x^{j_{p+1}}} \qquad (3.15)$$

in (3.12). Under these circumstances (3.13) becomes

$$\mathcal{O}(d\omega) = -\sum_{r=1}^{p} (-1)^{r-1} \left\{ \int_0^1 t^p \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^j} dt \right\} x^{j_r} dx^j$$
$$\wedge dx^{j_1} \wedge \cdots \wedge dx^{j_{r-1}} \wedge dx^{j_{r+1}} \wedge \cdots \wedge dx^{j_p}$$

$$+ \left\{ \int_0^1 t^p \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^j} dt \right\} x^j dx^{j_1} \wedge \cdots \wedge dx^{j_p}. \qquad (3.16)$$

Thus, when (3.11) and (3.16) are added, we obtain

$$d(\mathcal{O}\omega) + \mathcal{O}(d\omega)$$
$$= \int_0^1 \left[ t^p \frac{\partial A_{j_1 \cdots j_p}(tx^h)}{\partial x^j} x^j + pt^{p-1} A_{j_1 \cdots j_p}(tx^h) \right] dt \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}$$
$$= \int_0^1 \frac{d}{dt} \left[ t^p A_{j_1 \cdots j_p}(tx^h) \right] dt \, dx^{j_1} \wedge \cdots \wedge dx^{j_p}$$
$$= A_{j_1 \cdots j_p}(x^h) \, dx^{j_1} \wedge \cdots \wedge dx^{j_p} = \omega. \qquad (3.17)$$

This is an identity which is valid for any $p$-form $\omega$ whose coefficients are of class $C^1$ on $U$; in particular, if $d\omega = 0$ on $U$, it follows from (3.17) that

$$\omega = d\pi, \qquad (3.18)$$

where the $(p-1)$-form $\pi$ is given by (3.10):

$$\pi = \mathcal{O}\omega. \qquad (3.19)$$

Thus $\omega$ is exact on $U$. This establishes the following theorem.

### CONVERSE OF POINCARÉ'S LEMMA

*Let $U$ be an open region of $X_n$ which is star-shaped with respect to a point $P$ on $X_n$. Then every closed $p$-form whose coefficients are defined and of class $C^1$ on $U$ is exact at $P$.*

*Remark.* The above method of proof not only establishes the existence of the required $(p-1)$-form $\pi$, but also, according to (3.19), indicates a method whereby $\pi$ may be evaluated as a $(p-1)$-form with class $C^1$ coefficients on $U$. However, it is immediately obvious that (3.19) does not represent a unique solution. For, let $\mu$ be any $(p-2)$-form, by means of which we can define the $(p-1)$-form

$$\tilde{\pi} = \pi + d\mu. \qquad (3.20)$$

Clearly $d\tilde{\pi} = d\pi$, so that (3.20) represents a class of $(p-1)$-forms satisfying the condition $\omega = d\tilde{\pi}$.

We shall now consider some extremely important special cases of the above theorem. Let us suppose that the 1-form $\omega = A_j \, dx^j$ is closed, that is, such that $d\omega = 0$. From (3.3) and (3.4) we infer that this is equivalent to the statement that

$$\frac{\partial A_j}{\partial x^k} - \frac{\partial A_k}{\partial x^j} = 0. \qquad (3.21)$$

However, the converse of Poincaré's Lemma now assures us of the existence of a 0-form $\phi$ [i.e., a scalar function $\phi(x^h)$], for which $\omega = d\phi$, which implies that (3.6) is satisfied. It follows that *the system* (3.21) *always ensures the existence of a function $\phi$ satisfying* (3.6). Since (3.6) obviously also implies (3.21), we conclude: *in order that a covariant vector field be a gradient field, it is necessary and sufficient that its curl vanishes.*

Now let us turn to a 2-form

$$\omega = A_{jh}\, dx^j \wedge dx^h, \tag{3.22}$$

for which

$$d\omega = \frac{\partial A_{jh}}{\partial x^k}\, dx^j \wedge dx^h \wedge dx^k. \tag{3.23}$$

If this form is closed, it follows from the remark at the end of Section 5.1 that

$$\delta^{rst}_{jhk} \frac{\partial A_{rs}}{\partial x^t} = 0,$$

or

$$\left(\frac{\partial A_{jh}}{\partial x^k} - \frac{\partial A_{jk}}{\partial x^h}\right) + \left(\frac{\partial A_{hk}}{\partial x^j} - \frac{\partial A_{hj}}{\partial x^k}\right) + \left(\frac{\partial A_{kj}}{\partial x^h} - \frac{\partial A_{kh}}{\partial x^j}\right) = 0. \tag{3.24}$$

From the converse of Poincaré's Lemma we infer the existence of a 1-form $\pi = B_h\, dx^h$ such that $\omega = d\pi$, or

$$\left(A_{jh} - \frac{\partial B_h}{\partial x^j}\right) dx^j \wedge dx^h = 0,$$

which in turn implies that

$$\delta^{rs}_{jh}\left(A_{rs} - \frac{\partial B_s}{\partial x^r}\right) = 0,$$

or

$$A_{jh} - A_{hj} = \frac{\partial B_h}{\partial x^j} - \frac{\partial B_j}{\partial x^h}. \tag{3.25}$$

*Hence the system* (3.24) *ensures the existence of functions $B_h$ such that* (3.25) *is satisfied.* In particular, if $A_{hj}$ is skew-symmetric, the system (3.24) reduces to

$$\frac{\partial A_{jh}}{\partial x^k} + \frac{\partial A_{hk}}{\partial x^j} + \frac{\partial A_{kj}}{\partial x^h} = 0, \tag{3.26}$$

while (3.25) becomes

$$A_{jh} = \frac{1}{2}\left(\frac{\partial B_h}{\partial x^j} - \frac{\partial B_j}{\partial x^h}\right). \tag{3.27}$$

Conversely, it follows by direct differentiation that (3.27) implies (3.26). Thus, *in order that a skew-symmetric tensor $A_{jh}$ be the curl of a vector field, it is necessary and sufficient that it satisfies the system* (3.26). [In this connection we recall that, for a type (0, 2) tensor field $A_{jh}$, the left-hand side of (3.26) is a type (0, 3) tensor field, so that the invariance of the "integrability conditions" (3.26) is guaranteed.]

Finally, turning to the general case of a $p$-form $\omega$ such as (3.9), we observe that the condition $d\omega = 0$ yields the system of equations

$$\delta^{h_1 \cdots h_{p+1}}_{j_1 \cdots j_{p+1}} \frac{\partial A_{h_1 \cdots h_p}}{\partial x^{h_{p+1}}} = 0 \tag{3.28}$$

on the one hand, while on the other the converse to Poincaré's Lemma ensures the existence of a $(p-1)$-form

$$\pi = B_{j_1 \cdots j_{p-1}}\, dx^{j_1} \wedge \cdots \wedge dx^{j_{p-1}}, \tag{3.29}$$

such that $\omega = d\pi$, or

$$\left[A_{j_1 \cdots j_p} - (-1)^{p-1} \frac{\partial B_{j_1 \cdots j_{p-1}}}{\partial x^{j_p}}\right] dx^{j_1} \wedge \cdots \wedge dx^{j_p} = 0, \tag{3.30}$$

which is equivalent to

$$\delta^{h_1 \cdots h_p}_{j_1 \cdots j_p}\left[A_{h_1 \cdots h_p} - (-1)^{p-1} \frac{\partial B_{h_1 \cdots h_{p-1}}}{\partial x^{h_p}}\right] = 0. \tag{3.31}$$

It is therefore inferred that *the system* (3.28) *ensures the existence of the functions $B_{h_1 \cdots h_{p-1}}$ such that* (3.31) *is satisfied.* Again, if $A_{j_1 \cdots j_p}$ represents a type (0, $p$) tensor, the "integrability" conditions (3.28) of (3.31) are tensor equations.

In the application of these integrability conditions the local nature of the solutions, as specified by the converse of Poincaré's Lemma, must naturally be observed.

## 5.4  SYSTEMS OF TOTAL DIFFERENTIAL EQUATIONS

In many applications of the exterior differential calculus one encounters systems of total differential equations which are described by vanishing 1-forms of a type somewhat more general than those encountered above. These will be briefly described in the present section, and to this end we shall

consider $n$ variables $x^j$, together with a set of $m$ variables $t^\alpha$, which we shall regard as the local coordinates of a region $G$ of an $(n + m)$-dimensional differentiable manifold $X_{n+m}$. [In this section Latin indices range from 1 to $n$ as before, while Greek indices range from 1 to $m$; the summation convention is operative in respect of both sets of indices.]

An equation of the type $x^j = x^j(t^\alpha)$ describes an $m$-dimensional subspace $C_m$ of $X_{n+m}$: for instance, if $m = 1$, this subspace is a curve. More generally, for a set of $n$ independent parameters $v^h$, the system of equations

$$x^j = x^j(t^\alpha, v^h), \qquad (4.1)$$

represents an $n$-parameter family of $m$-dimensional subspaces $C_m(v^h)$ of $X_{n+m}$. We shall assume that the functions (4.1) are of class $C^2$ in $t^\alpha$ and of class $C^1$ in $v^h$, and that the family covers the region $G$ of $X_{n+m}$ simply in the sense that through each point of $G$ there passes one member of $C_m(v^h)$ of the family. The latter assumption is tantamount to the requirement that the $n$ equations (4.1) can be solved for the $n$ parameters $v^h$:

$$v^h = g^h(t^\alpha, x^j). \qquad (4.2)$$

Let $P$ be a point with coordinates $(t^\alpha, x^j)$ on the subspace $C_m(v^h)$. It follows from (4.1) that a displacement $(dt^\alpha, dx^j)$ at $P$ tangential to $C_m(v^h)$ must satisfy the conditions

$$dx^j = b_\alpha^j \, dt^\alpha, \qquad (4.3)$$

in which

$$b_\alpha^j = \frac{\partial x^j(t^\varepsilon, v^h)}{\partial t^\alpha}, \qquad (4.4)$$

where it is understood that the parameters $v^h$ assume the values prescribed by (4.2) at the point $P$. In this sense the $nm$ functions (4.4) are uniquely defined as class $C^1$ functions of $(t^\varepsilon, x^h)$ on $G$ by the system (4.1):

$$b_\alpha^j = b_\alpha^j(t^\varepsilon, x^h). \qquad (4.5)$$

Moreover, these functions are tensors of types (1, 0) and (0, 1) with respect to coordinate transformations $\bar{x}^j = \bar{x}^j(x^h)$ and $\bar{t}^\alpha = \bar{t}^\alpha(t^\varepsilon)$, respectively.

Conversely, let us suppose that we are given a class $C^1$ tensor field (4.5) of this type on $G$. This allows us to define the 1-forms

$$\omega^j = dx^j - b_\alpha^j(t^\varepsilon, x^h) \, dt^\alpha \qquad \text{\textbullet} \qquad (4.6)$$

in $x^j$ and $t^\alpha$, of which it will be assumed that they are linearly independent. The system of $n$ total differential equations

$$\omega^j = 0 \qquad (4.7)$$

is said to be *completely integrable* if there exist $n$ class $C^2$ functions $x^j = x^j(t^\alpha, v^h)$ which are such that the substitution of these functions in (4.6) yields (4.7). If $m > 1$ this is possible if and only if the $b_\alpha^j$ satisfy certain conditions, for clearly the integrability of (4.7) implies that the solutions $x^j(t^\alpha, v^h)$ must satisfy

$$\frac{\partial x^j}{\partial t^\alpha} = b_\alpha^j(t^\varepsilon, x^h), \qquad (4.8)$$

which represents a system of $mn$ partial differential equations for the $n$ functions $x^j(t^\varepsilon, x^h)$. Thus, when $m > 1$, the system (4.8) is overdetermined in the sense that there are more equations than unknowns.

A necessary condition for the compatability of the members of the system (4.8) is easily obtained as follows. For, if we are given a class $C^2$ solution $x^j(t^\alpha, v^h)$ of (4.8), we obtain by differentiation of (4.8) with respect to $t^\beta$ and subsequent substitution from (4.8):

$$\frac{\partial^2 x^j}{\partial t^\beta \, \partial t^\alpha} = \frac{\partial b_\alpha^j}{\partial t^\beta} + \frac{\partial b_\alpha^j}{\partial x^h} \frac{\partial x^h}{\partial t^\beta} = \frac{\partial b_\alpha^j}{\partial t^\beta} + \frac{\partial b_\alpha^j}{\partial x^h} b_\beta^h,$$

which must be symmetric in $\alpha$, $\beta$. It is therefore necessary that

$$Q_\beta{}^j{}_\alpha = 0, \qquad (4.9)$$

where we have written

$$Q_\beta{}^j{}_\alpha = \frac{\partial b_\alpha^j}{\partial t^\beta} - \frac{\partial b_\beta^j}{\partial t^\alpha} + \frac{\partial b_\alpha^j}{\partial x^h} b_\beta^h - \frac{\partial b_\beta^j}{\partial x^h} b_\alpha^h. \qquad (4.10)$$

It will be seen that the condition (4.9) is also sufficient for the complete integrability of the system (4.7). Indeed, we shall now prove the following theorem, which is often associated with the name of Frobenius.

**THEOREM**

*In order that the system (4.7) of $n$ total differential equations with class $C^1$ coefficients $b_\alpha^j(t^\varepsilon, x^h)$ be completely integrable, it is necessary and sufficient that these coefficients satisfy the conditions (4.9).*

*Proof.* The necessity of (4.9) has already been established. In order to prove sufficiency, we consider some fixed point with coordinates $(t_0^\alpha, x_0^j)$ in $G$, and for any point with coordinates $(t^\alpha, x^j)$ in $G$ we write

$$t^\alpha = t_0^\alpha + \lambda^\alpha s \qquad (4.11)$$

with suitably chosen parameters $\lambda^\alpha$, $s$. Let us now consider the system of $n$ ordinary differential equations

$$\frac{dx^j}{ds} = \lambda^\alpha b_\alpha^j(t_0^\varepsilon + \lambda^\varepsilon s, x^h). \tag{4.12}$$

According to the general theory of such equations we are assured of the existence of class $C^2$ solutions of (4.12) of the type

$$x^j = \phi^j(s, \lambda^\varepsilon, v^h), \tag{4.13}$$

which are uniquely specified in terms of their parameters $v^h$ provided that suitable initial conditions are prescribed. The latter are here taken to be

$$x_0^j = \phi^j(0, \lambda^\varepsilon, v^h) = v^j, \tag{4.14}$$

the parameters $v^j$ having been chosen such that $v^j = x_0^j$ for $t^\alpha = t_0^\alpha$. These conditions also imply that

$$x_0^j = \phi^j(s, 0, v^h) = v^j, \tag{4.15}$$

since the solutions of (4.12) with $\lambda^\alpha = 0$ are constant. By definition, the functions $\phi^j$ satisfy

$$\frac{\partial \phi^j}{\partial s} = \lambda^\alpha b_\alpha^j(t_0^\varepsilon + \lambda^\varepsilon s, \phi^h). \tag{4.16}$$

We now introduce a new set of quantities $w_\alpha^j$ by writing

$$w_\alpha^j(s, \lambda^\varepsilon, v^h) = \frac{\partial \phi^j}{\partial \lambda^\alpha} - s b_\alpha^j(t_0^\varepsilon + \lambda^\varepsilon s, \phi^h). \tag{4.17}$$

Let us differentiate (4.16) with respect to $\lambda^\alpha$, after which we substitute from (4.17):

$$\frac{\partial^2 \phi^j}{\partial \lambda^\alpha \partial s} = b_\alpha^j + s \lambda^\beta \frac{\partial b_\beta^j}{\partial t^\alpha} + \lambda^\beta \frac{\partial b_\beta^j}{\partial x^h} \frac{\partial \phi^h}{\partial \lambda^\alpha}$$

$$= b_\alpha^j + s \lambda^\beta \left( \frac{\partial b_\beta^j}{\partial t^\alpha} + \frac{\partial b_\beta^j}{\partial x^h} b_\alpha^h \right) + \lambda^\beta \frac{\partial b_\beta^j}{\partial x^h} w_\alpha^h. \tag{4.18}$$

Similarly we differentiate (4.17) with respect to $s$, after which we substitute from (4.16):

$$\frac{\partial w_\alpha^j}{\partial s} = \frac{\partial^2 \phi^j}{\partial s \partial \lambda^\alpha} - b_\alpha^j - s \lambda^\beta \frac{\partial b_\alpha^j}{\partial t^\beta} - s \frac{\partial b_\alpha^j}{\partial x^h} \frac{\partial \phi^h}{\partial s}$$

$$= \frac{\partial^2 \phi^j}{\partial s \partial \lambda^\alpha} - b_\alpha^j - s \lambda^\beta \left( \frac{\partial b_\alpha^j}{\partial t^\beta} + \frac{\partial b_\alpha^j}{\partial x^h} b_\beta^h \right). \tag{4.19}$$

This is added to (4.18), it being recalled that $\phi^j$ is of class $C^2$. In terms of the notation (4.10) we thus obtain

$$\frac{\partial w_\alpha^j}{\partial s} - \lambda^\beta \frac{\partial b_\beta^j}{\partial x^h} w_\alpha^h = s \lambda^\beta Q_{\alpha\ \beta}^{\ j}. \tag{4.20}$$

At this stage we invoke the assumption (4.9), so that (4.20) reduces to

$$\frac{\partial w_\alpha^j}{\partial s} - \lambda^\beta \frac{\partial b_\beta^j}{\partial x^h} w_\alpha^h = 0. \tag{4.21}$$

For fixed values of the parameters $\lambda^\alpha$, $v^h$, we can regard this as a system of ordinary linear differential equations for $w_\alpha^j$ as functions of $s$. But from (4.17) we infer that $w_\alpha^j = 0$ for $s = 0$, since by virtue of (4.14) the derivatives $\partial \phi^j / \partial \lambda^\alpha$ vanish when $s = 0$. Thus, from the general theory of linear systems such as (4.21), it follows that

$$w_\alpha^j = 0 \tag{4.22}$$

for all values of $s$. Hence (4.17) reduces to

$$\frac{\partial \phi^j}{\partial \lambda^\alpha} = s b_\alpha^j(t_0^\varepsilon + \lambda^\varepsilon s, \phi^h). \tag{4.23}$$

Using (4.11) we now construct the following functions for $s \neq 0$:

$$X^j(s, t^\alpha, v^h) = \phi^j(s, \lambda^\alpha, v^h) = \phi^j \left( s, \frac{t^\alpha - t_0^\alpha}{s}, v^h \right). \tag{4.24}$$

On differentiating this with respect to $t^\alpha$, taking into account (4.23) and (4.11), we obtain

$$\frac{\partial X^j}{\partial t^\alpha} = \frac{\partial \phi^j}{\partial \lambda^\beta} \frac{1}{s} \delta_\alpha^\beta = \frac{1}{s} \frac{\partial \phi^j}{\partial \lambda^\alpha} = b_\alpha^j(t^\varepsilon, \phi^h). \tag{4.25}$$

Also, differentiation of (4.24) with respect to $s$ yields

$$\frac{\partial X^j}{\partial s} = \frac{\partial \phi^j}{\partial s} - \frac{\partial \phi^j}{\partial \lambda^\alpha} \frac{t^\alpha - t_0^\alpha}{s^2} = \frac{\partial \phi^j}{\partial s} - \frac{\partial \phi^j}{\partial \lambda^\alpha} \frac{\lambda^\alpha}{s},$$

or, by virtue of (4.16) and (4.23),

$$\frac{\partial X^j}{\partial s} = \lambda^\alpha b_\alpha^j(t^\varepsilon, \phi^h) - \lambda^\alpha b_\alpha^j(t^\varepsilon, \phi^h) = 0, \tag{4.26}$$

from which it follows that the functions (4.24) do not contain $s$ explicitly. We may therefore write

$$x^j(t^\alpha, v^h) = X^j(s, t^\alpha, v^h) = \phi^j(s, \lambda^\alpha, v^h), \tag{4.27}$$

and according to (4.25) we then have

$$\frac{\partial x^j}{\partial t^\alpha} = b_\alpha^j(t^\varepsilon, x^h),\tag{4.28}$$

which is (4.8). Finally, putting $\lambda^\alpha = 0$ in (4.27), noting that this implies $t^\alpha = t_0^\alpha$ by (4.11), we find, using (4.15),

$$x^j(t_0^\alpha, v^h) = \phi^j(s, 0, v^h) = x_0^j.\tag{4.29}$$

From (4.28) and (4.29) we infer that the functions $x^j(t^\alpha, v^h)$ constructed according to (4.27) are solutions of the system (4.7) and represent an $m$-dimensional surface in $X_{m+n}$ passing through the point $P$ with coordinates $(t_0^\alpha, x_0^j)$. By variation of the parameters $v^h$ an $n$-parameter family (4.1) of such surfaces is obtained. This concludes the proof of the theorem.

The latter possesses an important corollary. Having established the existence of the family (4.1) as a consequence of the condition (4.9), we may construct the corresponding functions (4.2), which, by definition, are integrals of the system (4.7) in the sense that these functions are constant on each member $C_m(v^h)$ of the family (4.1). Thus, for an arbitrary displacement $dt^\alpha$, we have

$$0 = \frac{\partial g^j}{\partial t^\alpha} dt^\alpha + \frac{\partial g^j}{\partial x^h} dx^h = \left(\frac{\partial g^j}{\partial t^\alpha} + b_\alpha^h \frac{\partial g^j}{\partial x^h}\right) dt^\alpha,$$

from which we deduce that each of the $n$ independent functions (4.2) is a solution of the partial differential equation

$$\frac{\partial \Phi(t^\varepsilon, x^k)}{\partial t^\alpha} + b_\alpha^h(t^\varepsilon, x^k)\frac{\partial \Phi(t^\varepsilon, x^k)}{\partial x^h} = 0.\tag{4.30}$$

It may therefore be inferred that *the conditions* (4.9) *ensure the existence of n functionally independent solutions* $g^j(t^\varepsilon, x^k)$ *of* (4.30). Moreover, if $F$ denotes any class $C^1$ function of $n$ independent variables, it follows directly that $F(g^j(t^\varepsilon, x^k))$ *is also a solution of* (4.30), for

$$\frac{\partial F}{\partial t^\alpha} + b_\alpha^h \frac{\partial F}{\partial x^h} = \frac{\partial F}{\partial g^j}\frac{\partial g^j}{\partial t^\alpha} + b_\alpha^h \frac{\partial F}{\partial g^j}\frac{\partial g^j}{\partial x^h} = \frac{\partial F}{\partial g^j}\left(\frac{\partial g^j}{\partial t^\alpha} + b_\alpha^h \frac{\partial g^j}{\partial x^h}\right) = 0.$$

Again, *any solution of* (4.30) *is an integral of* (4.7).

We shall now show how the theorem of Frobenius can be translated into the language of the exterior differential calculus. By virtue of (2.1) the exterior derivative of the 1-form (4.6) is given by

$$d\omega^j = -\frac{\partial b_\alpha^j}{\partial t^\beta} dt^\beta \wedge dt^\alpha - \frac{\partial b_\alpha^j}{\partial x^h} dx^h \wedge dt^\alpha,$$

or, if we eliminate $dx^h$ by means of (4.6),

$$d\omega^j = \left(\frac{\partial b_\alpha^j}{\partial t^\beta} + \frac{\partial b_\alpha^j}{\partial x^h} b_\beta^h\right) dt^\alpha \wedge dt^\beta + \frac{\partial b_\alpha^j}{\partial x^h} dt^\alpha \wedge \omega^h.$$

In terms of the notation (4.10) this can be written in the form

$$d\omega^j = \sum_{\alpha < \beta} Q_{\beta\,\alpha}^{\ j} dt^\alpha \wedge dt^\beta + \frac{\partial b_\alpha^j}{\partial x^h} dt^\alpha \wedge \omega^h.\tag{4.31}$$

Let us now put

$$\Omega^j = d\omega^j \wedge \omega^1 \wedge \cdots \wedge \omega^n.\tag{4.32}$$

An explicit expression for this $(n + 2)$-form is obtained by taking the exterior product of (4.31) with $\omega^1 \wedge \cdots \wedge \omega^n$, where it is to be noted that, because of (1.3),

$$\omega^h \wedge \omega^1 \wedge \cdots \wedge \omega^n = 0 \qquad (h = 1, \ldots, n).$$

It is thus found that

$$\Omega^j = \sum_{\alpha < \beta} Q_{\beta\,\alpha}^{\ j} dt^\alpha \wedge dt^\beta \wedge \omega^1 \wedge \cdots \wedge \omega^n.\tag{4.33}$$

It follows that the condition (4.9) implies that $\Omega^j = 0$. In order to show that the converse is also true, we note that the 1-forms with which we have been dealing here [such as (4.6)] are in the variables $(x^j, t^\alpha)$, so that the basis elements of the vector space of these 1-forms can be chosen to be $(dx^1, dx^2, \ldots, dx^n, dt^1, \ldots, dt^m)$. The $(n + m) \times (n + m)$ matrix

$$\begin{pmatrix} \delta_h^j & -b_\alpha^j \\ 0 & \delta_\beta^\alpha \end{pmatrix},\tag{4.34}$$

whose determinant has the value $+1$, when applied to these basis elements yields a new basis, namely $(\omega^1, \ldots, \omega^n, dt^1, \ldots, dt^m)$. Thus the $(n + 2)$-forms

$$\omega^1 \wedge \cdots \wedge \omega^n \wedge dt^\alpha \wedge dt^\beta, \qquad \text{with } \alpha < \beta,$$

constitute a subset of the corresponding basis elements of the $\binom{n+m}{n+2}$-dimensional vector space of all $(n + 2)$-forms. It therefore follows from (4.33) that $\Omega^j = 0$ implies that $Q_{\beta\,\alpha}^{\ j} = 0$: consequently the conditions $\Omega^j = 0$ and (4.9) are equivalent. The theorem of Frobenius may therefore be enunciated in the following manner: *In order that the system* (4.7) *of n total differential equations with class* $C^1$ *coefficients* $b_\alpha^j(t^\varepsilon, x^h)$ *be completely integrable it is necessary and sufficient that the* $(n + 2)$-*forms* $\Omega^j$ *as defined by* (4.32) *vanish*:

$$d\omega^j \wedge \omega^1 \wedge \cdots \wedge \omega^n = 0 \qquad (j = 1, \ldots, n).\tag{4.35}$$

Because of the structure of the 1-forms (4.6), our system (4.7) of total differential equations is of a somewhat special kind. However, the above formulation of the theorem of Frobenius is easily extended to more general systems (which is one of the principal advantages of this formulation). Let us therefore consider the system of $n$ total differential equations described by $\pi^j = 0$, where the 1-forms $\pi^j$ are defined by

$$\pi^j = G_h^j(t^\varepsilon, x^k)\, dx^h + G_\alpha^j(t^\varepsilon, x^k)\, dt^\alpha, \tag{4.36}$$

it being assumed that the coefficients in (4.36) are of class $C^1$ and such that

$$\det(G_h^j) \neq 0. \tag{4.37}$$

Under these circumstances the inverse $H_j^k(t^\varepsilon, x^k)$ of $G_h^j(t^\varepsilon, x^k)$ exists:

$$H_j^k G_h^j = \delta_h^k. \tag{4.38}$$

One can therefore define the following $n$ 1-forms:

$$\mu^k = H_j^k \pi^j = dx^k - B_\alpha^k(t^\varepsilon, x^h)\, dt^\alpha, \tag{4.39}$$

where

$$B_\alpha^k = -H_j^k G_\alpha^j. \tag{4.40}$$

Clearly the 1-forms (4.39) possess the structure of the 1-forms (4.6), and thus the system

$$\mu^k = 0 \tag{4.41}$$

is integrable if and only if the $(n + 2)$-forms

$$U^k = d\mu^k \wedge \mu^1 \wedge \cdots \wedge \mu^n \qquad (k = 1, \ldots, n) \tag{4.42}$$

vanish. But, by virtue of the construction (4.39), the integrability of (4.41) implies that of the given system $\pi^j = 0$, and it remains to rewrite this criterion in terms of the given 1-forms (4.36).

Now, from (4.39) we deduce that

$$d\mu^k = dH_j^k \wedge \pi^j + H_j^k\, d\pi^j,$$

or, since the inverse of (4.39) is

$$\pi^j = G_h^j \mu^h \tag{4.43}$$

in view of (4.38), we have

$$d\mu^k = G_h^j(dH_j^k \wedge \mu^h) + H_j^k\, d\pi^j. \tag{4.44}$$

Again, on taking the exterior product of (4.44) with $\mu^1 \wedge \cdots \wedge \mu^n$, noting that $\mu^h \wedge \mu^1 \wedge \cdots \wedge \mu^n = 0$ for $h = 1, \ldots, n$, we find that the $(n + 2)$-forms

(4.42) are given by

$$U^k = H_j^k\, d\pi^j \wedge \mu^1 \wedge \cdots \wedge \mu^n. \tag{4.45}$$

But from (4.39) we have

$$\mu^1 \wedge \cdots \wedge \mu^n = H_{j_1}^1 \cdots H_{j_n}^n \pi^{j_1} \wedge \cdots \wedge \pi^{j_n}$$

$$= \frac{1}{n!} \delta_{h_1 \cdots h_n}^{j_1 \cdots j_n} H_{j_1}^1 \cdots H_{j_n}^n \pi^{h_1} \wedge \cdots \wedge \pi^{h_n} = \det(H_h^l)\pi^1 \wedge \cdots \wedge \pi^n,$$

and accordingly (4.45) becomes

$$U^k = \det(H_h^l)H_j^k\, d\pi^j \wedge \pi^1 \wedge \cdots \wedge \pi^n. \tag{4.46}$$

By virtue of the assumption (4.37) it therefore follows that the condition $U^k = 0$ is equivalent to

$$\Pi^j = 0, \tag{4.47}$$

where

$$\Pi^j = d\pi^j \wedge \pi^1 \wedge \cdots \wedge \pi^n. \tag{4.48}$$

It may thus be inferred that the system $\pi^j = 0$ of $n$ total differential equations is integrable if and only if the conditions (4.47) are satisfied.

In order to facilitate a direct application of this result to general systems of total differential equations it may be advisable to change the notation slightly. Let us consider a set of $N$ independent variables $z^A$. [Here, and for the remainder of this section, capital Latin indices $A$, $B$, $C$ range from 1 to $N$; the summation convention applies to these indices also.] It is supposed that a system of $n$ total differential equations of the form

$$\omega^j = 0 \qquad (j = 1, \ldots, n; n < N) \tag{4.49}$$

is given, where

$$\omega^j = G_A^j(z^C)\, dz^A. \tag{4.50}$$

From the above theory it now follows that *the system* (4.49) *is integrable if and only if the $n$ conditions*

$$d\omega^j \wedge \omega^1 \wedge \cdots \wedge \omega^n = 0 \qquad (j = 1, \ldots, n) \tag{4.51}$$

*are satisfied.*

As an example, let us consider the case when $n = 1$. The given total differential equation is then of the form

$$\omega = G_A\, dz^A = 0.$$

Clearly

$$d\omega \wedge \omega = \frac{\partial G_A}{\partial z^B}\, dz^B \wedge dz^A \wedge \omega = -G_C \frac{\partial G_A}{\partial z^B}\, dz^A \wedge dz^B \wedge dz^C.$$

Thus, in view of the concluding remark of Section 5.1, the integrability condition $d\omega \wedge \omega = 0$ is equivalent to the requirement that

$$G_A\left(\frac{\partial G_B}{\partial z^C} - \frac{\partial G_C}{\partial z^B}\right) + G_B\left(\frac{\partial G_C}{\partial z^A} - \frac{\partial G_A}{\partial z^C}\right) + G_C\left(\frac{\partial G_A}{\partial z^B} - \frac{\partial G_B}{\partial z^A}\right) = 0.$$

When $N = 3$ this reduces to a single relation of the type $\mathbf{G} \cdot \text{curl } \mathbf{G} = 0$, which is a necessary and sufficient condition that $\mathbf{G}$ be proportional to the gradient of a scalar function.

## 5.5  THE THEOREM OF STOKES

One of the most powerful tools of the calculus of exterior differential forms is the so-called theorem of Stokes, which consists of a very general integral formula. Indeed, the latter contains all the known integral theorems of classical vector analysis, such as the divergence theorem of Gauss, as special cases. However, whereas the formulation of the divergence theorem depends crucially on the Euclidean (or Riemannian) metric of the underlying manifold, the formula of Stokes is entirely independent of any metrical concepts.

A rigorous proof of Stokes' theorem in its fullest generality involves concepts and techniques from algebraic topology which are beyond the scope of this text. Accordingly a somewhat less ambitious treatment is presented here, as a result of which the final formula does not reflect Stokes' theorem in its fullest generality, particularly as regards the topological nature of admissible domains of integration. Nevertheless, the result to be derived below should suffice for most purposes.

Since the theorem of Stokes is concerned with the integration of $p$-forms over $p$-dimensional regions, it is necessary to begin with the definition of the integral of a $p$-form. Let $D$ denote an $n$-dimensional measurable region of our differentiable manifold $X_n$, and let us suppose that we are given $n + 1$ scalar functions $f, \phi^1, \ldots, \phi^n$ which are of class $C^1$ in their arguments $x^h$ on $D$. We may therefore write, by virtue of (1.17),

$$d\phi^1 \wedge \cdots \wedge d\phi^n = \frac{\partial(\phi^1, \ldots, \phi^n)}{\partial(x^1, \ldots, x^n)} dx^1 \wedge \cdots \wedge dx^n. \tag{5.1}$$

Accordingly we now formulate the following definition†:

$$\int_D f \, d\phi^1 \wedge \cdots \wedge d\phi^n = \int_D f(x^h) \frac{\partial(\phi^1, \ldots, \phi^n)}{\partial(x^1, \ldots, x^n)} dx^1 \cdots dx^n, \tag{5.2}$$

where the right-hand side denotes an $n$-fold integral in the usual sense.

† A rigorous and comprehensive definition of the process of integration of differential forms on manifolds is beyond the scope of this text. See, for instance, Bishop and Goldberg [1].

*The value of this integral is independent of the choice of the coordinate system,* for, under a coordinate transformation $\bar{x}^j = \bar{x}^j(x^h)$, with positive Jacobian,

$$\frac{\partial \phi^r}{\partial \bar{x}^j} = \frac{\partial \phi^r}{\partial x^h}\frac{\partial x^h}{\partial \bar{x}^j},$$

so that

$$\frac{\partial(\phi^1, \ldots, \phi^n)}{\partial(\bar{x}^1, \ldots, \bar{x}^n)} = \frac{\partial(\phi^1, \ldots, \phi^n)}{\partial(x^1, \ldots, x^n)} \frac{\partial(x^1, \ldots, x^n)}{\partial(\bar{x}^1, \ldots, \bar{x}^n)}. \tag{5.3}$$

Proceeding in a similar fashion to subspaces of $X_n$, let us consider a measurable region $B$ of a $p$-dimensional subspace, which is represented parametrically by class $C^1$ functions of the form

$$x^j = x^j(u^\alpha) \qquad (\alpha = 1, \ldots, p) \tag{5.4}$$

in which the $u^\alpha$ denote the $p$ parameters. For any given set of $p$ class $C^1$ functions $\psi^\alpha(x^h)$ we define by analogy with (5.2):

$$\int_B f \, d\psi^1 \wedge \cdots \wedge d\psi^p = \int_B f(x^h(u^\alpha)) \frac{\partial(\psi^1(x^h(u^\alpha)), \ldots, \psi^p(x^h(u^\alpha)))}{\partial(u^1, \qquad . \qquad . \qquad . \qquad, u^p)} du^1 \cdots du^p, \tag{5.5}$$

where it is to be understood that no summation over $p$ is implied. As above, it is easily shown that the value of this integral is independent of the choice of the parametrization (5.4) of the region $B$.

We shall now turn to the derivation of Stokes' theorem. Let $G$ be a $(p + 1)$-dimensional region of $X_n$ which is star-shaped relative to some point $O$ in the interior of $G$ and bounded by a closed $p$-dimensional subspace $\partial G$. It is assumed that the latter is covered by a finite number of coordinate neighborhoods $U_1, U_2, \ldots$, such that the corresponding parametrizations are of class $C^1$. If $u^1, \ldots, u^p$ denote a set of parameters corresponding to $U_1$, a region $B$ of $\partial G$ is determined by a closed set $b$ in the domain $R_p$ of the variables $u^1, \ldots, u^p$. The region $B$ is represented parametrically by the equations

$$\xi^j = \xi^j(u^\alpha), \qquad u^\alpha \in b \qquad (\alpha = 1, \ldots, p) \tag{5.6}$$

where the $\xi^j$ refer to a special coordinate system of $X_n$ whose origin is located at $O$. By hypothesis, the set of points

$$C(B) = \{t\xi^j(u^\alpha): \quad 0 \le t \le 1, \quad u^\alpha \in b\}$$

defines a $(p + 1)$-dimensional region of $X_n$ which is contained entirely in $G$.

Let $f^A(\xi^h)$ denote $(p + 1)$ differentiable functions of which it is assumed that they are homogeneous of degree 1 in their arguments:

$$f^A(t\xi^h) = tf^A(\xi^h) \quad \text{for all } t \ge 0 \qquad (A = 0, 1, \ldots, p). \tag{5.7}$$

Since the $(p + 1)$ variables $t$, $u^\alpha$ may obviously be used to define a parametrization of $C(B)$, we have, according to the definition (5.5):

$$\int_{C(B)} df^0 \wedge df^1 \wedge \cdots \wedge df^p$$

$$= \int_{C(B)} \frac{\partial(f^0(t\xi^h(u^\alpha)), f^1(t\xi^h(u^\alpha)), \ldots, f^p(t\xi(u^\alpha)))}{\partial(t, u^1, \quad . \quad . \quad . \quad , u^p)} \, dt \, du^1 \cdots du^p. \tag{5.8}$$

Let us now put

$$g^A(u^\beta) = f^A(\xi^h(u^\beta)) \qquad (A = 0, 1, \ldots, p; \ \beta = 1, \ldots, p) \tag{5.9}$$

so that, by (5.7),

$$f^A(t\xi^h(u^\beta)) = tg^A(u^\beta),$$

from which it follows that

$$\frac{\partial f^A(t\xi^h(u^\beta))}{\partial t} = g^A(u^\beta), \qquad \frac{\partial f^A(t\xi^h(u^\beta))}{\partial u^\alpha} = t \frac{\partial g^A}{\partial u^\beta}.$$

Thus the functional determinant in the integrand of (5.8) becomes

$$t^p \begin{vmatrix} g^0 & g^1 & \cdots & g^p \\ \dfrac{\partial g^0}{\partial u^1} & \dfrac{\partial g^1}{\partial u^1} & \cdots & \dfrac{\partial g^p}{\partial u^1} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial g^0}{\partial u^p} & \dfrac{\partial g^1}{\partial u^p} & \cdots & \dfrac{\partial g^p}{\partial u^p} \end{vmatrix} = t^p \sum_{A=0}^{p} (-1)^A g^A \frac{\partial(g^0, g^1, \ldots, g^{A-1}, g^{A+1}, \ldots, g^p)}{\partial(u^1, \quad . \quad . \quad . \quad , u^p)}$$

after expansion in terms of the elements of the first row. Except for the factor $t^p$, this expression depends solely on the $p$ variables $u^\beta$, and thus (5.8) can be written in the form

$$\int_{C(B)} df^0 \wedge df^1 \wedge \cdots \wedge df^p$$

$$= \sum_{A=0}^{p} (-1)^A \int_0^1 t^p \, dt \int_B g^A \frac{\partial(g^0, g^1, \ldots, g^{A-1}, g^{A+1}, \ldots, g^p)}{\partial(u^1, \quad . \quad . \quad . \quad , u^p)} \, du^1 \cdots du^p$$

$$= \frac{1}{p+1} \sum_{A=0}^{p} (-1)^A \int_B g^A \, dg^0 \wedge dg^1 \wedge \cdots \wedge dg^{A-1} \wedge dg^{A+1} \wedge \cdots \wedge dg^p, \tag{5.10}$$

where, in the last step, we have used the general formula (5.5).

The relation (5.10) is an integral formula which holds for any set of $(p + 1)$ functions $f^A(\xi^h)$ satisfying condition (5.7). The latter will hold if, in particular,

we put

$$f^A(\xi^h) = \xi^{A+1} = g^A \qquad (A = 0, \ldots, p) \tag{5.11}$$

so that (5.10) reduces to

$$\int_{C(B)} d\xi^1 \wedge \cdots \wedge d\xi^{p+1} = \frac{1}{p+1} \sum_{A=1}^{p+1} (-1)^{A+1} \int_B \xi^A \, d\xi^1 \wedge \cdots \wedge d\xi^{A-1}$$

$$\wedge \, d\xi^{A+1} \wedge \cdots \wedge d\xi^{p+1}. \tag{5.12}$$

By hypothesis, the boundary $\partial G$ of $G$ can be considered as the union of regions $B_1, B_2, \ldots$, which correspond to the coordinate neighborhoods $U_1, U_2, \ldots$, respectively, such that none of $B_1, B_2, \ldots$ have interior points in common. The given star-shaped region $G$ then consists of the union of $C(B_1), C(B_2), \ldots$, for each of which a formula of the type (5.12) holds. On adding the latter, we obtain

$$\int_G d\xi^1 \wedge \cdots \wedge d\xi^{p+1} = \frac{1}{p+1} \sum_{A=1}^{p+1} (-1)^{A+1} \int_{\partial G} \xi^A \, d\xi^1 \wedge \cdots \wedge d\xi^{A-1}$$

$$\wedge \, d\xi^{A+1} \wedge \cdots \wedge d\xi^{p+1}. \tag{5.13}$$

Now let us introduce a one-parameter coordinate transformation $x^j = x^j(\xi^h, \lambda)$ with nonvanishing Jacobian, which is such that the inverse transformation may be expressed in the form

$$\xi^j = x^j + \lambda \phi^j(x^h), \tag{5.14}$$

where the functions $\phi^j(x^h)$ are of class $C^1$ on $G$ but otherwise quite arbitrary. In passing we note that

$$\frac{\partial \xi^j}{\partial x^h} = \delta_h^j + \lambda \frac{\partial \phi^j}{\partial x^h}, \tag{5.15}$$

so that, for sufficiently small values of $\lambda$, the Jacobian of (5.14) is in fact positive on $G$. When (5.14) is substituted in (5.13), the left-hand side of the latter becomes

$$\int_G (dx^1 + \lambda \, d\phi^1) \wedge \cdots \wedge (dx^{p+1} + \lambda \, d\phi^{p+1}),$$

in which the coefficient of $\lambda^{p+1}$ is

$$\int_G d\phi^1 \wedge \cdots \wedge d\phi^{p+1}. \tag{5.16}$$

On the other hand, the right-hand side of (5.13) becomes, after substitution from (5.14),

$$\frac{1}{p+1} \sum_{A=1}^{p+1} (-1)^{A+1} \int_{\partial G} (x^A + \lambda \phi^A)(dx^1 + \lambda \, d\phi^1) \wedge \cdots \wedge (dx^{A-1} + \lambda \, d\phi^{A-1})$$

$$\wedge (dx^{A+1} + \lambda \, d\phi^{A+1}) \wedge \cdots \wedge (dx^{p+1} + \lambda \, d\phi^{p+1}),$$

in which the coefficient of $\lambda^{p+1}$ is

$$\frac{1}{p+1} \sum_{A=1}^{p+1} (-1)^{A+1} \int_{\partial G} \phi^A \, d\phi^1 \wedge \cdots \wedge d\phi^{A-1} \wedge d\phi^{A+1} \wedge \cdots \wedge d\phi^{p+1}.$$

$$(5.17)$$

Thus a comparison of (5.16) with (5.17) yields

$$(p+1) \int_G d\phi^1 \wedge \cdots \wedge d\phi^{p+1} = \sum_{A=1}^{p+1} (-1)^{A+1} \int_{\partial G} \phi^A \, d\phi^1 \wedge \cdots \wedge d\phi^{A-1}$$

$$\wedge d\phi^{A+1} \wedge \cdots \wedge d\phi^{p+1}.$$

$$(5.18)$$

This integral formula holds for any set of $(p+1)$ class $C^1$ functions $\phi^A(x^j)$. In particular, with $\phi^{p+1} = 1$, we have $d\phi^{p+1} = 0$, and the only surviving term in (5.18) occurs on the right-hand side, namely, the term corresponding to the value $A = p + 1$. It is therefore inferred that

$$\int_{\partial G} d\phi^1 \wedge \cdots \wedge d\phi^p = 0 \qquad (5.19)$$

for *any* set of $p$ functions $\phi^1, \ldots, \phi^p$.

Returning to the more general formula (5.18), let us consider, for example, the first two terms on the right-hand side, which are

$$\int_{\partial G} \phi^1 \, d\phi^2 \wedge d\phi^3 \wedge \cdots \wedge d\phi^{p+1} \qquad (5.20)$$

and

$$-\int_{\partial G} \phi^2 \, d\phi^1 \wedge d\phi^3 \wedge \cdots \wedge d\phi^{p+1} \qquad (5.21)$$

respectively. Since $\phi^1$, $\phi^2$ are 0-forms, the expression (5.21) can be written in the form

$$\int_{\partial G} \phi^1 \, d\phi^2 \wedge d\phi^3 \wedge \cdots \wedge d\phi^{p+1} - \int_{\partial G} d(\phi^1 \phi^2) \wedge d\phi^3 \wedge \cdots \wedge d\phi^{p+1}.$$

$$(5.22)$$

However, it follows from (5.19) that the second integral on the right-hand side of (5.22) vanishes identically, and accordingly a comparison of (5.20) with (5.22) indicates that the first two terms of the sum in (5.18) are identical. By symmetry, it is obvious that all terms of this sum have the same value, and, since there are $p + 1$ such terms, it follows that (5.18) is equivalent to

$$\int_G d\phi^1 \wedge \cdots \wedge d\phi^{p+1} = \int_{\partial G} \phi^1 \, d\phi^2 \wedge d\phi^3 \wedge \cdots \wedge d\phi^{p+1}. \quad (5.23)$$

Clearly this integral formula is valid for any set of $(p + 1)$ functions $\phi^1, \ldots, \phi^{p+1}$ of class $C^1$.

Let us now define a $p$-form $\omega$ by putting

$$\omega = \phi^1 \, d\phi^2 \wedge d\phi^3 \wedge \cdots \wedge d\phi^{p+1}, \qquad (5.24)$$

so that

$$d\omega = d\phi^1 \wedge d\phi^2 \wedge \cdots \wedge d\phi^{p+1}. \qquad (5.25)$$

Moreover, by suitable choice of $\phi^1, \ldots, \phi^{p+1}$ *any* $p$-form whose coefficients are continuous functions of $x^h$ can be expressed in the form (5.24), or as linear combinations of such forms. Substitution of (5.24) and (5.25) in (5.23) therefore yields the following theorem.

## STOKES' THEOREM

*Let $G$ be a $(p + 1)$-dimensional region of $X_n$, star-shaped with respect to an interior point $O$, the boundary $\partial G$ of $G$ being a closed $p$-dimensional subspace admitting parametric representations of class $C^1$. Then, for any $p$-form $\omega$ defined on $G$ with continuously differentiable coefficients,*

$$\int_G d\omega = \int_{\partial G} \omega. \qquad (5.26)$$

We shall now consider a few special cases of this theorem. Let us construct an $(n - 1)$-form $\pi_j$ defined by

$$(n-1)! \, \pi_j = \varepsilon_{j j_2 \cdots j_n} \, dx^{j_2} \wedge \cdots \wedge dx^{j_n}, \qquad (5.27)$$

so that

$$\pi_j = (-1)^{j+1} \, dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n. \qquad (5.28)$$

It should be recalled that the permutation symbol $\varepsilon_{j_1 \cdots j_n}$ is a type $(0, n)$ relative tensor of weight $-1$ (Section 4.2), so that $\pi_j$ is a covariant relative vector of weight $-1$. Now let $A^j(x^h)$ denote a differentiable contravariant vector field of weight $+1$ on $G$: then the $(n - 1)$-form $\omega$ defined by

$$\omega = \pi_j A^j \qquad (5.29)$$

is clearly a scalar. From the definition (5.28) it follows that

$$d\pi_j = 0 \tag{5.30}$$

identically, and hence (5.29) yields

$$d\omega = \frac{\partial A^j}{\partial x^h} dx^h \wedge \pi_j$$

$$= (-1)^{j+1} \frac{\partial A^j}{\partial x^h} dx^h \wedge dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n,$$

in which a summation over both $h$ and $j$ is implied. For a given value of $j$, the only surviving term in the summation corresponding to the index $h$ is that for which $h = j$; collecting all terms of this kind, we obtain

$$d\omega = \frac{\partial A^j}{\partial x^j} dx^1 \wedge \cdots \wedge dx^n. \tag{5.31}$$

Since $A^j$ represents a contravariant relative vector of weight $+1$, the sum $\partial A^j / \partial x^j$ is tensorial; in fact, according to (4.1.27) this expression is identical with $A^j_{|j}$ (whenever $X_n$ is endowed with a connection). Hence (5.31) is equivalent to

$$d\omega = A^j_{|j} dx^1 \wedge \cdots \wedge dx^n. \tag{5.32}$$

Thus the substitution of (5.29) and (5.32) in Stokes' formula (5.26) with $p = n - 1$ yields

$$\int_G \frac{\partial A^j}{\partial x^j} dx^1 \wedge \cdots \wedge dx^n = \int_G A^j_{|j} dx^1 \wedge \cdots \wedge dx^n = \int_{\partial G} \pi_j A^j, \tag{5.33}$$

*which is a divergence-type integral theorem*, since the integrand of the $n$-fold integrals on the left is the divergence of the field $A^j$. Moreover, if the parametric representation of the $(n-1)$-dimensional hypersurface is given by (5.4), with $\alpha = 1, \ldots, n - 1$, a displacement $dx^j$ tangential to $\partial G$ may be represented in the form

$$dx^j = \frac{\partial x^j}{\partial u^\alpha} du^\alpha; \tag{5.34}$$

thus on $\partial G$ the $(n-1)$-form $\pi_j$ as defined by (5.27) can be expressed as

$$(n-1)! \pi_j = \varepsilon_{jj_2\cdots j_n} \frac{\partial x^{j_2}}{\partial u^{\alpha_1}} \cdots \frac{\partial x^{j_n}}{\partial u^{\alpha_{n-1}}} du^{\alpha_1} \wedge \cdots \wedge du^{\alpha_{n-1}}$$

$$= \varepsilon_{jj_2\cdots j_n} \frac{\partial x^{j_2}}{\partial u^{\alpha_1}} \cdots \frac{\partial x^{j_n}}{\partial u^{\alpha_{n-1}}} \varepsilon^{\alpha_1\cdots\alpha_{n-1}} du^1 \wedge \cdots \wedge du^{n-1}$$

$$= \varepsilon_{jj_2\cdots j_n} \frac{\partial(x^{j_2}, \ldots, x^{j_n})}{\partial(u^1, \ldots, u^{n-1})} du^1 \wedge \cdots \wedge du^{n-1}. \tag{5.35}$$

However, it should be observed that

$$\varepsilon_{jj_2\cdots j_n} \frac{\partial x^j}{\partial u^\alpha} \frac{\partial(x^{j_2}, \ldots, x^{j_n})}{\partial(u^1, \ldots, u^{n-1})} = \frac{\partial(x^j, x^{j_2}, \ldots, x^{j_n})}{\partial(u^\alpha, u^1, \ldots, u^{n-1})} = 0 \tag{5.36}$$

identically for each value of $\alpha$ ($\alpha = 1, \ldots, n - 1$) since the second determinant in (5.36) contains two identical rows. It therefore follows from (5.35) and (5.36) that

$$\pi_j \frac{\partial x^j}{\partial u^\alpha} = 0 \qquad (\alpha = 1, \ldots, n - 1) \tag{5.37}$$

identically. This result may be interpreted geometrically in the following manner. The $(n-1)$ vectors $\partial x^j / \partial u^1, \ldots, \partial x^j / \partial u^{n-1}$ at any point $P$ of $\partial G$ are tangent to $\partial G$ at $P$ by construction; in other words, they span the tangent hyperplane of $G$ in the tangent space $T_n(P)$ of $P$. The left-hand side of (5.37) is an inner product: if, in analogy with metric geometry, the vanishing of the inner product is interpreted as orthogonality, one may interpret (5.37) by the statement that *the relative vector field $\pi_j$ is orthogonal (or normal) to $\partial G$*. This result clearly illuminates the similarity between the nonmetrical divergence theorem (5.33) and the classical divergence theorem of Gauss.

As a further example, let us consider the explicit form of (5.26) on a two-dimensional Euclidean plane referred to Cartesian coordinates $x^1$, $x^2$. Any 1-form on $E_2$ is of the form $\omega = A_j dx^j$ ($j = 1, 2$), so that

$$d\omega = \frac{\partial A_j}{\partial x^h} dx^h \wedge dx^j = \left( \frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} \right) dx^1 \wedge dx^2.$$

For a region $G$ of $E_2$ bounded by a closed, simple curve $C$, we then have from (5.26) with $p = 1$:

$$\int_G \left( \frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} \right) dx^1 \wedge dx^2 = \oint_C \omega = \oint_C (A_1 dx^1 + A_2 dx^2), \tag{5.38}$$

which is often referred to as Green's theorem in the plane.

Finally, let us put $p = 0$ in (5.26), so that $G$ is a continuous curve $C$ which is bounded by the initial and final points $P_1(t = t_1)$ and $P_2(t = t_2)$, respectively. For any differentiable 0-form $\omega = f(t)$, with $d\omega = f'(t) dt$, the formula (5.26) gives

$$\int_{t_1}^{t_2} f'(t) dt = f(t_2) - f(t_1),$$

which shows that Stokes' theorem may be regarded as a generalization of the fundamental theorem of the calculus.

## 5.6 CURVATURE FORMS ON DIFFERENTIABLE MANIFOLDS

We shall now return to our $n$-dimensional differentiable manifold $X_n$ of which it is once more assumed that it is endowed with a connection, the latter not being necessarily symmetric. It will be shown how the curvature theory of $X_n$ as presented in Sections 3.6–3.8 may be rederived by means of an application of the formal theory of exterior differential forms on $X_n$.

Let $\Pi^j$ be some $p$-form on $X_n$, this form being defined by a class $C^2$ type $(1, p)$ tensor field on $X_n$. According to (2.25) its covariant exterior derivative is given by

$$D\Pi^j = d\Pi^j + \omega_h{}^j \wedge \Pi^h, \tag{6.1}$$

which is a $(p + 1)$-form of the same type, whose exterior covariant derivative is therefore

$$D(D\Pi^j) = d(D\Pi^j) + \omega_l{}^j \wedge D\Pi^l.$$

In order to find an explicit expression for this $(p + 2)$-form, we substitute from (6.1), after which (2.7) and (2.8) are applied, which gives

$$
\begin{aligned}
D(D\Pi^j) &= d(d\Pi^j + \omega_h{}^j \wedge \Pi^h) + \omega_l{}^j \wedge (d\Pi^l + \omega_h{}^l \wedge \Pi^h) \\
&= d\omega_h{}^j \wedge \Pi^h - \omega_h{}^j \wedge d\Pi^h + \omega_l{}^j \wedge d\Pi^l + \omega_l{}^j \wedge \omega_h{}^l \wedge \Pi^h \\
&= (d\omega_h{}^j + \omega_l{}^j \wedge \omega_h{}^l) \wedge \Pi^h,
\end{aligned}
$$

or

$$D(D\Pi^j) = \Omega_h{}^j \wedge \Pi^h, \tag{6.2}$$

where we have put

$$\Omega_h{}^j = d\omega_h{}^j + \omega_l{}^j \wedge \omega_h{}^l. \tag{6.3}$$

The form thus defined is called the *curvature 2-form*, for, if we substitute from (2.20) in (6.3), we obtain

$$
\begin{aligned}
\Omega_h{}^j &= \frac{\partial \Gamma_h{}^j{}_k}{\partial x^l} dx^l \wedge dx^k + \Gamma_m{}^j{}_k \Gamma_h{}^m{}_l \, dx^k \wedge dx^l \\
&= -\left( \frac{\partial \Gamma_h{}^j{}_k}{\partial x^l} - \Gamma_m{}^j{}_k \Gamma_h{}^m{}_l \right) dx^k \wedge dx^l \\
&= -\frac{1}{2} \left( \frac{\partial \Gamma_h{}^j{}_k}{\partial x^l} - \frac{\partial \Gamma_h{}^j{}_l}{\partial x^k} + \Gamma_m{}^j{}_l \Gamma_h{}^m{}_k - \Gamma_m{}^j{}_k \Gamma_h{}^m{}_l \right) dx^k \wedge dx^l.
\end{aligned}
$$

Clearly the coefficient of $-\frac{1}{2} dx^k \wedge dx^l$ on the right-hand side is the curvature tensor $K_h{}^j{}_{kl}$ as defined by (3.6.8), so that

$$\Omega_l{}^j = -\tfrac{1}{2} K_l{}^j{}_{hk} \, dx^h \wedge dx^k, \tag{6.4}$$

from which it is evident that the curvature 2-form is a type $(1, 1)$ tensor.

Closely associated with this form is the *torsion 2-form* which is defined by

$$\Omega^j = dx^h \wedge \omega_h{}^j. \tag{6.5}$$

Again it follows directly from (2.20) that this form may be expressed as

$$\Omega^j = \Gamma_h{}^j{}_k \, dx^h \wedge dx^k = \tfrac{1}{2}(\Gamma_h{}^j{}_k - \Gamma_k{}^j{}_h) \, dx^h \wedge dx^k,$$

or, in terms of the torsion tensor (2.15),

$$\Omega^j = \tfrac{1}{2} S_h{}^j{}_k \, dx^h \wedge dx^k. \tag{6.6}$$

Thus the torsion 2-form is a type $(1, 0)$ tensor. Moreover, if we apply (6.1) to the 1-forms $dx^j$, noting (2.8), (1.23), and (6.5), we infer that

$$D(dx^j) = d(dx^j) + \omega_h{}^j \wedge dx^h = -dx^h \wedge \omega_h{}^j = -\Omega^j. \tag{6.7}$$

It is remarkable that $D(dx^j) = 0$ identically whenever the connection is symmetric.

For the special case $\Pi^j = dx^j$ in (6.2), the latter becomes

$$D[D(dx^j)] = \Omega_h{}^j \wedge dx^h, \tag{6.8}$$

and this, because of (6.7), yields

$$D\Omega^j = -\Omega_h{}^j \wedge dx^h, \tag{6.9}$$

which clearly indicates how the covariant exterior derivative of the torsion 2-form is related to the curvature 2-form.

Now, by virtue of the definition (2.27) we have

$$D\Omega_h{}^j = d\Omega_h{}^j + \omega_l{}^j \wedge \Omega_h{}^l - \omega_h{}^l \wedge \Omega_l{}^j. \tag{6.10}$$

But, because of (2.7) and (2.8), it follows directly from (6.3) that

$$d\Omega_h{}^j = d\omega_l{}^j \wedge \omega_h{}^l - \omega_l{}^j \wedge d\omega_h{}^l,$$

in which we substitute from (6.3), which gives

$$
\begin{aligned}
d\Omega_h{}^j &= (\Omega_l{}^j - \omega_m{}^j \wedge \omega_l{}^m) \wedge \omega_h{}^l - \omega_l{}^j \wedge (\Omega_h{}^l - \omega_m{}^l \wedge \omega_h{}^m) \\
&= \omega_h{}^l \wedge \Omega_l{}^j - \omega_l{}^j \wedge \Omega_h{}^l,
\end{aligned} \tag{6.11}
$$

since the two terms involving the exterior products of the three 1-forms $\omega_l{}^m$ cancel with each other. From (6.10) one therefore obtains *the following remarkable identity*:

$$D\Omega_h{}^j = 0. \tag{6.12}$$

The relations (6.3) and (6.5) are called the *equations of structure* (in the terminology of E. Cartan); clearly the identities (6.9) and (6.12) are direct consequences of these equations. We shall now interpret the above formalism in terms of the theory of Sections 3.6–3.8.

Recalling the remarks concerning the 1-form (2.30), let us regard the absolute differential $DX^j$ of a type $(1, 0)$ tensor field $X^j$ as a 1-form, namely as $DX^j = X^j_{|h} \, dx^h$. By virtue of the product rule (2.31) we have

$$D(DX^j) = D(X^j_{|h} \, dx^h) = (DX^j_{|h}) \wedge dx^h + X^j_{|h} \, D(dx^h),$$

or, if we apply (6.7),

$$D(DX^j) = X^j_{|h|k} \, dx^k \wedge dx^h - \Omega^h X^j_{|h}. \tag{6.13}$$

However, with $\Pi^j = X^j$ in the relation (6.2), the latter becomes

$$D(DX^j) = \Omega_h{}^j X^h,$$

so that (6.13) yields

$$-X^j_{|h|k} \, dx^h \wedge dx^k - \Omega^h X^j_{|h} = \Omega_h{}^j X^h.$$

To this relation we now apply (6.4) and (6.6), which gives

$$X^j_{|h|k} \, dx^h \wedge dx^k = \tfrac{1}{2} K_l{}^j{}_{hk} X^l \, dx^h \wedge dx^k - \tfrac{1}{2} S_h{}^l{}_k X^j_{|l} \, dx^h \wedge dx^k,$$

or, by virtue of the skew-symmetry of the curvature and torsion tensors in the subscripts $h$ and $k$,

$$\sum_{h<k} (X^j_{|h|k} - X^j_{|k|h} - K_l{}^j{}_{hk} X^l + S_h{}^l{}_k X^j_{|l}) \, dx^h \wedge dx^k = 0, \tag{6.14}$$

which is obviously equivalent to the Ricci identity (3.6.9). More generally, then, *the relation* (6.2) *implies the Ricci identities for type* $(1, p)$ *tensors.*

Let us now turn to the identity (6.9). Because of (6.6), we have by analogy with (2.23):

$$D\Omega^j = \tfrac{1}{2}(S_h{}^j{}_{k|l} + S_m{}^j{}_k S_h{}^m{}_l) \, dx^h \wedge dx^k \wedge dx^l. \tag{6.15}$$

When this, together with (6.4), is substituted in (6.9), we obtain

$$(S_h{}^j{}_{k|l} + S_m{}^j{}_k S_h{}^m{}_l - K_l{}^j{}_{hk}) \, dx^h \wedge dx^k \wedge dx^l = 0. \tag{6.16}$$

According to the concluding remarks of Section 5.1, this implies that

$$\delta^{rst}_{hkl}(S_r{}^j{}_{s|t} + S_m{}^j{}_s S_r{}^m{}_t - K_t{}^j{}_{rs}) = 0,$$

which, because of the skew-symmetry properties of the curvature and torsion tensors, is equivalent to the cyclic identity (3.8.6). Furthermore, for the curvature 2-form (6.4) we have, by virtue of (2.29),

$$D\Omega_j{}^l = -\tfrac{1}{2}(K_j{}^l{}_{hk|p} + K_j{}^l{}_{mk} S_h{}^m{}_p) \, dx^h \wedge dx^k \wedge dx^p. \tag{6.17}$$

Thus the identity (6.12) implies that

$$\delta^{rst}_{hkp}(K_j{}^l{}_{rs|t} + K_j{}^l{}_{ms} S_r{}^m{}_t) = 0, \tag{6.18}$$

which obviously gives rise to the Bianchi identity (3.8.11). *Thus the identities* (6.9) *and* (6.12) *are equivalent to the cyclic and Bianchi identities, respectively.*

Finally, let us turn to the problem of the parallel displacement of a vector field $X^j$ along a curve of $X_n$ as specified by the conditions (3.7.5), which, in the notation (2.30), can be expressed in the form

$$\omega^j = 0, \tag{6.19}$$

where

$$\omega^j = dX^j + \omega_h{}^j X^h. \tag{6.20}$$

The system (6.19) of $n$ total differential equations is of the type exemplified by (4.6), for in terms of (2.20) we can write (6.20) in the form

$$\omega^j = dX^j - b_k^j(x^l, X^l) \, dx^k, \tag{6.21}$$

where

$$b_k^j(x^l, X^l) = -\Gamma_m{}^j{}_k X^m. \tag{6.22}$$

A parallel vector field over a finite $n$-dimensional region of $X_n$ is defined if and only if the system (6.19) is completely integrable. According to the theorem of Frobenius this is possible if and only if the conditions

$$Q_k{}^j{}_h = 0 \tag{6.23}$$

are satisfied, where, by (4.10),

$$Q_k{}^j{}_h = \frac{\partial b_h^j}{\partial x^k} - \frac{\partial b_k^j}{\partial x^h} + \frac{\partial b_h^j}{\partial X^l} b_k^l - \frac{\partial b_k^j}{\partial X^l} b_h^l. \tag{6.24}$$

But from (6.22) we obviously have

$$\frac{\partial b_h^j}{\partial x^k} = -\frac{\partial \Gamma_m{}^j{}_h}{\partial x^k} X^m, \qquad \frac{\partial b_h^j}{\partial X^l} = -\Gamma_l{}^j{}_h,$$

and thus (6.24) gives

$$Q_k{}^j{}_h = \left( -\frac{\partial \Gamma_m{}^j{}_h}{\partial x^k} + \frac{\partial \Gamma_m{}^j{}_k}{\partial x^h} + \Gamma_l{}^j{}_h \Gamma_m{}^l{}_k - \Gamma_l{}^j{}_k \Gamma_m{}^l{}_h \right) X^m,$$

or, in terms of the definition (3.6.8),

$$Q_k{}^j{}_h = -K_m{}^j{}_{hk} X^m = K_m{}^j{}_{kh} X^m. \tag{6.25}$$

From (6.23) it is therefore inferred that *the necessary and sufficient condition for the existence of a parallel vector field* $X^j$ *on a finite region of* $X_n$ *is represented by the equations*

$$K_m{}^j{}_{hk} X^m = 0. \tag{6.26}$$

This, of course, is in agreement with the conclusion (3.7.15) of Section 3.7; however, in that section merely the necessity, and not the sufficiency, of (6.26) is established.

Alternatively, the condition (4.35) could have been applied to the system (6.19). From (6.20) we have, using (2.7) and (2.8),

$$d\omega^j = d\omega_h{}^j X^h - \omega_h{}^j \wedge dX^h,$$

or, after substitution from (6.20),

$$d\omega^j = d\omega_h{}^j X^h - \omega_h{}^j \wedge \omega^h + \omega_l{}^j \wedge \omega_h{}^l X^h.$$

On observing (6.3) we therefore infer that

$$d\omega^j = \Omega_h{}^j X^h - \omega_h{}^j \wedge \omega^h. \tag{6.27}$$

It follows that

$$d\omega^j \wedge \omega^1 \wedge \cdots \wedge \omega^n = (\Omega_h{}^j \wedge \omega^1 \wedge \cdots \wedge \omega^n)X^h. \tag{6.28}$$

Thus the necessary and sufficient conditions (4.35) for the complete integrability of the system (6.19) are satisfied if

$$\Omega_h{}^j X^h = 0, \tag{6.29}$$

which, according to (6.4), is equivalent to (6.26). Moreover, an argument similar to that following (4.33) involving the construction of a subset of appropriate basis elements of $(n + 2)$-forms in $(x^h, X^h)$ may be used to show that (6.29) also implies (4.35).

Let us now suppose that the curvature tensor vanishes everywhere on $X_n$. A given set of $n$ linearly independent contravariant vectors $X^j_{(h)}$ at a point $P_0$ of $X_n$ then defines, by parallel displacement, a set of $n$ vector fields $X^j_{(h)}(x^l)$ on $X_n$ which satisfy the system of partial differential equations

$$\frac{\partial X^j_{(h)}}{\partial x^l} = -\Gamma_m{}^j{}_l X^m_{(h)}, \tag{6.30}$$

the integrability of which is guaranteed by our assumption. (Here the subscript $h$ enclosed in parentheses is a mere labeling index which distinguishes the $n$ distinct vector fields.) From the transformation law for the contravariant vectors $X^j_{(h)}$ it follows that $X = \det(X^j_{(h)})$ is a relative scalar of weight $-1$. Thus, if $Y^{(h)}_j$ denotes the cofactor of $X^j_{(h)}$ in $X$, we have

$$X_{|k} = \frac{\partial X}{\partial x^k} + X\Gamma_l{}^l{}_k = \sum_{h=1}^{n} Y^{(h)}_j \frac{\partial X^j_{(h)}}{\partial x^k} + X\Gamma_j{}^j{}_k,$$

or, if we use (6.30),

$$X_{|k} = -\sum_{h=1}^{n} Y^{(h)}_j \Gamma_m{}^j{}_k X^m_{(h)} + X\Gamma_j{}^j{}_k = -\delta^m_j X\Gamma_m{}^j{}_k + X\Gamma_j{}^j{}_k = 0. \tag{6.31}$$

However, at the initial point $P_0$ we have $X \neq 0$ by virtue of the linear independence of the $X^j_{(h)}$ at $P_0$, and thus (6.31) indicates that $X \neq 0$ throughout $X_n$.

We shall now endeavor to introduce a new coordinate system $\bar{x}^h = \bar{x}^h(x^j)$ by means of the partial differential equations

$$\frac{\partial x^j}{\partial \bar{x}^h} = X^j_{(h)}(x^l). \tag{6.32}$$

According to (4.10) the integrability conditions for this system are given by

$$\frac{\partial X^j_{(h)}}{\partial x^l} X^l_{(k)} - \frac{\partial X^j_{(k)}}{\partial x^l} X^l_{(h)} = 0,$$

or, if we use (6.30),

$$\Gamma_m{}^j{}_l X^m_{(h)} X^l_{(k)} - \Gamma_m{}^j{}_l X^m_{(k)} X^l_{(h)} = (\Gamma_m{}^j{}_l - \Gamma_l{}^j{}_m)X^m_{(h)} X^l_{(k)} = 0.$$

Since the $X^j_{(h)}$ at $P_0$ are essentially arbitrary, this can be satisfied if and only if our connection is symmetric, as will now be assumed. Moreover, according to the remarks made above, the Jacobian $J = \det(X^j_{(h)})$ of (6.32) is nonvanishing throughout $X_n$, and thus (6.32) defines an admissible coordinate transformation on $X_n$. Because of (6.30) and (6.32) the second derivatives of this transformation are given by

$$\frac{\partial^2 x^j}{\partial \bar{x}^h \partial \bar{x}^k} = \frac{\partial X^j_{(h)}}{\partial x^l} \frac{\partial x^l}{\partial \bar{x}^k} = -\Gamma_m{}^j{}_l X^m_{(h)} \frac{\partial x^l}{\partial \bar{x}^k} = -\Gamma_m{}^j{}_l \frac{\partial x^m}{\partial \bar{x}^h} \frac{\partial x^l}{\partial \bar{x}^k}. \tag{6.33}$$

But, by virtue of (3.3.21), the connection coefficients in the $\bar{x}$-coordinate system are given quite generally by

$$\bar{\Gamma}_h{}^p{}_k = \frac{\partial x^p}{\partial \bar{x}^j}\left(\Gamma_m{}^j{}_l \frac{\partial x^m}{\partial \bar{x}^h} \frac{\partial x^l}{\partial \bar{x}^k} + \frac{\partial^2 x^j}{\partial \bar{x}^h \partial \bar{x}^k}\right), \tag{6.34}$$

and substitution of (6.33) in (6.34) yields $\bar{\Gamma}_h{}^p{}_k = 0$. We have therefore established the following theorem.

**THEOREM**

*If the curvature tensor vanishes on a differentiable manifold $X_n$ endowed with a symmetric connection, there exist special coordinate systems relative to which the components of this connection vanish everywhere.*

## 5.7   SUBSPACES OF A DIFFERENTIABLE MANIFOLD

On our differentiable manifold $X_n$ a system of $n$ equations of the form

$$x^j = x^j(u^\alpha) \qquad (j = 1, \ldots, n; \alpha = 1, \ldots, m; m < n), \tag{7.1}$$

which express the $n$ local coordinates of a point $P$ of $X_n$ in terms of $m$ variables $u^\alpha$, represents an $m$-dimensional locus in $X_n$. [Throughout this section Greek indices range from 1 to $m$, and Latin indices continue to assume the values 1 to $n$; the summation convention is operative in respect of both sets of indices.] It will be assumed that the functions $x^j(u^\alpha)$ are of class $C^3$, and that the rank of the matrix of the first derivatives

$$B_\alpha^j = \frac{\partial x^j}{\partial u^\alpha} \tag{7.2}$$

is maximal, namely, $m$. The $m$-dimensional locus defined by (7.1) is clearly a differentiable manifold in its own right and is henceforth denoted by $X_m$. Since each point of $X_m$ is also a point of $X_n$, we say that $X_m$ is a *subspace* of $X_n$, or that $X_m$ is *embedded* (or immersed) in $X_n$. The system (7.1) is said to be a *parametric representation* of $X_m$, whose coordinates $u^\alpha$ are often called the *parameters* of $X_m$.

In the theory of subspaces one is confronted with two entirely distinct types of transformations. First, we have, as before, coordinate transformations on $X_n$ of the form

$$\bar{x}^j = \bar{x}^j(x^h), \tag{7.3}$$

while, second, we also have to consider parameter transformations

$$\tilde{u}^\alpha = \tilde{u}^\alpha(u^\beta) \tag{7.4}$$

on $X_m$. It is assumed henceforth that these transformations are of class $C^2$ with nonvanishing Jacobians; moreover, it should be emphasized that (7.3) and (7.4) are entirely independent of each other.

Putting $\bar{B}_\alpha^j = \partial \bar{x}^j / \partial u^\alpha$, we infer directly from (7.2) and (7.3) that

$$\bar{B}_\alpha^j = \frac{\partial \bar{x}^j}{\partial x^h} B_\alpha^h. \tag{7.5}$$

On the other hand, if we put $\tilde{B}_\alpha^j = \partial x^j / \partial \tilde{u}^\alpha$, it follows immediately from (7.2) and (7.4) that

$$\tilde{B}_\alpha^j = \frac{\partial u^\beta}{\partial \tilde{u}^\alpha} B_\beta^j. \tag{7.6}$$

Thus, relative to the coordinate transformation (7.3), the $B_\alpha^j$ behave as components of a type (1, 0) tensor, whereas relative to the parameter trans-

formation (7.4) the $B_\alpha^j$ behave as components of a type (0, 1) tensor. This phenomenon is typical of the theory of subspaces and provides the motivation for much of our subsequent analysis.

By virtue of (7.1) and (7.2) the $m$ 1-forms $du^1, \ldots, du^m$ at a point $P$ of $X_m$ determine $n$ 1-forms $dx^1, \ldots, dx^n$, given by

$$dx^j = B_\alpha^j \, du^\alpha, \tag{7.7}$$

which may be interpreted as the components of a displacement in $X_n$ which is tangential to $X_m$ at $P$. More generally, let $U^\alpha$ denote the components of an element of the tangent space $T_m(P)$ of $X_m$ at $P$, the $U^\alpha$ representing the components of a type (1, 0) tensor under the parameter transformation (7.4). By means of the relation

$$X^j = B_\alpha^j U^\alpha, \tag{7.8}$$

the components of an element of $T_n(P)$ are uniquely defined. Because of the linearity of (7.8) in $U^\alpha$ we may thus regard $T_m(P)$ as a linear subspace of $T_n(P)$ spanned by the $m$ linearly independent vectors $B_1^j, \ldots, B_m^j$, which, as we have seen, are themselves elements of $T_n(P)$. [If the embedding manifold $X_n$ were an $n$-dimensional Euclidean space $E_n$, the plane $T_m(P)$ would coincide with the tangent plane to $X_m$ in the usual sense.]

We shall now suppose that our manifolds $X_m$ and $X_n$ are each endowed with an affine connection, whose components will be denoted by $\Gamma_h{}^j{}_k$ and $\Gamma_\alpha{}^\varepsilon{}_\beta$, respectively. It is not assumed that these connections are symmetric, or that the $\Gamma_\alpha{}^\varepsilon{}_\beta$ are in any way related to the $\Gamma_h{}^j{}_k$. (Strictly speaking, one should not use the same symbol $\Gamma$ for the connection coefficients on both $X_m$ and $X_n$; however, since the former is always endowed with Greek indices, whereas the latter may only have Latin indices, the proposed notation should not cause any confusion.) These connections are introduced, in the first instance, merely in order to enable us to construct covariant derivatives of quantities defined on $X_m$ which are tensorial with respect to one or both of the transformations (7.3) and (7.4). By definition of connection, the transformation laws of $\Gamma_h{}^j{}_k$ and $\Gamma_\alpha{}^\varepsilon{}_\beta$ under (7.3) and (7.4) are respectively given by

$$\frac{\partial^2 \bar{x}^j}{\partial x^h \, \partial x^k} = \frac{\partial \bar{x}^j}{\partial x^l} \Gamma_h{}^l{}_k - \bar{\Gamma}_p{}^j{}_q \frac{\partial \bar{x}^p}{\partial x^h} \frac{\partial \bar{x}^q}{\partial x^k} \tag{7.9}$$

and

$$\frac{\partial^2 u^\varepsilon}{\partial \tilde{u}^\alpha \, \partial \tilde{u}^\beta} = \frac{\partial u^\varepsilon}{\partial \tilde{u}^\lambda} \tilde{\Gamma}_\alpha{}^\lambda{}_\beta - \Gamma_\sigma{}^\varepsilon{}_\gamma \frac{\partial u^\sigma}{\partial \tilde{u}^\alpha} \frac{\partial u^\gamma}{\partial \tilde{u}^\beta}. \tag{7.10}$$

In order to fix our ideas, let us consider a field $X_\alpha^j(u^\varepsilon)$ defined as class $C^1$ functions on $X_m$, representing the components of a type (1, 0) tensor relative to

to (7.3) and a type (0, 1) tensor relative to (7.4), so that

$$\bar{X}_\alpha^j = \frac{\partial \bar{x}^j}{\partial x^h} X_\alpha^h, \qquad \tilde{X}_\alpha^j = \frac{\partial u^\varepsilon}{\partial \tilde{u}^\alpha} X_\varepsilon^j. \tag{7.11}$$

It should be emphasized that $X_\alpha^j(u^\varepsilon)$ is given as a function of $u^1, \ldots, u^m$ (and not of $x^1, \ldots, x^n$), and accordingly only the derivatives $\partial X_\alpha^j/\partial u^\beta$ (and not $\partial X_\alpha^j/\partial x^h$) are defined. Differentiating the first and second members of (7.11) with respect to $u^\beta$ and $\tilde{u}^\beta$, respectively, and noting (7.2) we find

$$\frac{\partial \bar{X}_\alpha^j}{\partial u^\beta} = \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} B_\beta^k X_\alpha^h + \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial X_\alpha^h}{\partial u^\beta}, \tag{7.12}$$

and

$$\frac{\partial \tilde{X}_\alpha^j}{\partial \tilde{u}^\beta} = \frac{\partial^2 u^\varepsilon}{\partial \tilde{u}^\alpha \partial \tilde{u}^\beta} X_\varepsilon^j + \frac{\partial X_\varepsilon^j}{\partial u^\gamma} \frac{\partial u^\varepsilon}{\partial \tilde{u}^\alpha} \frac{\partial u^\gamma}{\partial \tilde{u}^\beta}. \tag{7.13}$$

Thus, as expected, neither of the derivatives (7.12) and (7.13) is tensorial in any way. In order to construct covariant derivatives according to the usual procedure, we eliminate the second derivatives in (7.12) by means of (7.9), which gives

$$\frac{\partial \bar{X}_\alpha^j}{\partial u^\beta} = \left( \frac{\partial \bar{x}^j}{\partial x^l} \Gamma_{h\,k}^l - \bar{\Gamma}_{p\,q}^j \frac{\partial \bar{x}^p}{\partial x^h} \frac{\partial \bar{x}^q}{\partial x^k} \right) B_\beta^k X_\alpha^h + \frac{\partial \bar{x}^j}{\partial x^l} \frac{\partial X_\alpha^l}{\partial u^\beta},$$

or, if we observe (7.5) and the first member of (7.11),

$$\frac{\partial \bar{X}_\alpha^j}{\partial u^\beta} + \bar{\Gamma}_{p\,q}^j \bar{X}_\alpha^p \bar{B}_\beta^q = \frac{\partial \bar{x}^j}{\partial x^l} \left( \frac{\partial X_\alpha^l}{\partial u^\beta} + \Gamma_{h\,k}^l X_\alpha^h B_\beta^k \right). \tag{7.14}$$

Similarly, an application of (7.10) to (7.13) yields

$$\frac{\partial \tilde{X}_\alpha^j}{\partial \tilde{u}^\beta} - \tilde{\Gamma}_{\alpha\,\beta}^\lambda \tilde{X}_\lambda^j = \frac{\partial u^\varepsilon}{\partial \tilde{u}^\alpha} \frac{\partial u^\gamma}{\partial \tilde{u}^\beta} \left( \frac{\partial X_\varepsilon^j}{\partial u^\gamma} - \Gamma_{\varepsilon\,\gamma}^\lambda X_\lambda^j \right). \tag{7.15}$$

It is therefore inferred that the quantities on the left-hand side of (7.14) represent the components of a type (1, 0) tensor under (7.3), while those on the left-hand side of (7.15) represent the components of a type (0, 2) tensor under (7.4). However, neither of these quantities is tensorial with respect to both of (7.3) and (7.4), and therefore they cannot be regarded as suitable covariant derivatives. Nevertheless, taken in conjunction, they suggest the following definition:

$$X_{\alpha\|\beta}^j = \frac{\partial X_\alpha^j}{\partial u^\beta} - \Gamma_{\alpha\,\beta}^\lambda X_\lambda^j + \Gamma_{h\,k}^j X_\alpha^h B_\beta^k. \tag{7.16}$$

These quantities do, in fact, represent the components of a tensor field which is of type (1, 0) relative to the coordinate transformation (7.3) and of type (0, 2) relative to the parameter transformation (7.4). This assertion follows directly from (7.14), (7.15) and the fact that the $\Gamma_{h\,k}^j$, $\Gamma_{\alpha\,\beta}^\lambda$ are invariant under (7.4) and (7.3), respectively. We shall therefore call (7.16) the *mixed covariant derivative* of the tensor field $X_\alpha^j(u^\varepsilon)$. Clearly this definition can be readily extended to tensors of arbitrary types relative to the transformations (7.3) and (7.4), or to arbitrary tensorial $p$-forms. The usual laws of covariant differentiation apply also to mixed covariant derivatives. However, it should be emphasized that, for quantities which are tensorial relative to (7.4), while being invariant relative to (7.3), the mixed covariant derivative is simply the usual covariant derivative with respect to the connection of $X_m$.

The *mixed absolute differential* corresponding to (7.16) is given by

$$DX_\alpha^j = X_{\alpha\|\beta}^j du^\beta = dX_\alpha^j - \Gamma_{\alpha\,\beta}^\lambda X_\lambda^j du^\beta + \Gamma_{h\,k}^j X_\alpha^h dx^k, \tag{7.17}$$

where, in the third term, we have used (7.7). If, as in (2.20), we put

$$\omega_\alpha^\lambda = \Gamma_{\alpha\,\beta}^\lambda du^\beta, \tag{7.18}$$

we can write (7.17) in the form

$$DX_\alpha^j = dX_\alpha^j - \omega_\alpha^\lambda X_\lambda^j + \omega_h^j X_\alpha^h. \tag{7.19}$$

By analogy with (6.5) we can introduce the torsion 2-form on $X_m$, which is defined by

$$\Omega^\alpha = du^\beta \wedge \omega_\beta^\alpha, \tag{7.20}$$

and for which

$$\Omega^\alpha = \tfrac{1}{2} S_{\varepsilon\,\gamma}^\alpha du^\varepsilon \wedge du^\gamma, \tag{7.21}$$

where

$$S_{\varepsilon\,\gamma}^\alpha = \Gamma_{\varepsilon\,\gamma}^\alpha - \Gamma_{\gamma\,\varepsilon}^\alpha \tag{7.22}$$

is the torsion tensor of $X_m$. As in (6.7) it follows that

$$D(du^\alpha) = -\Omega^\alpha. \tag{7.23}$$

Now let $\Pi_\alpha^j$ denote a set of $mn$ $p$-forms on $X_m$ $(0 \le p \le m)$ which may be represented in the form

$$\Pi_\alpha^j = A_{\alpha\beta_1\cdots\beta_p}^j du^{\beta_1} \wedge \cdots \wedge du^{\beta_p}, \tag{7.24}$$

where $A_{\alpha\beta_1\cdots\beta_p}^j$ is a type (1, 0) tensor relative to (7.3) and a type (0, $p + 1$) tensor relative to (7.4). Then

$$D\Pi_\alpha^j = d\Pi_\alpha^j - \omega_\alpha^\lambda \wedge \Pi_\lambda^j + \omega_h^j \wedge \Pi_\alpha^h \tag{7.25}$$

is a $(p + 1)$-form of the same type, for which

$$D(D\Pi_\alpha^j) = d(D\Pi_\alpha^j) - \omega_\alpha^\lambda \wedge (D\Pi_\lambda^j) + \omega_h^j \wedge (D\Pi_\alpha^h). \tag{7.26}$$

In order to evaluate this expression, we substitute from (7.25), carrying out the differentiation implied in the first term with the aid of (2.7). It is thus found that

$$\begin{aligned}
D(D\Pi_\alpha^j) = &-d\omega_\alpha^\lambda \wedge \Pi_\lambda^j + \omega_\alpha^\lambda \wedge d\Pi_\lambda^j \\
&+ d\omega_h^j \wedge \Pi_\alpha^h - \omega_h^j \wedge d\Pi_\alpha^h - \omega_\alpha^\lambda \wedge d\Pi_\lambda^j \\
&+ \omega_\alpha^\lambda \wedge \omega_\lambda^\sigma \wedge \Pi_\sigma^j - \omega_\alpha^\lambda \wedge \omega_h^j \wedge \Pi_\lambda^h \\
&+ \omega_h^j \wedge d\Pi_\alpha^h - \omega_h^j \wedge \omega_\alpha^\sigma \wedge \Pi_\sigma^h + \omega_h^j \wedge \omega_k^h \wedge \Pi_\alpha^k,
\end{aligned}$$

or, after some simplification with the use of (1.23),

$$D(D\Pi_\alpha^j) = \Omega_h^j \wedge \Pi_\alpha^h - \Omega_\alpha^\beta \wedge \Pi_\beta^j, \tag{7.27}$$

where the curvature 2-form $\Omega_h^j$ of $X_n$ is given by (6.3), while the *curvature 2-form* $\Omega_\alpha^\beta$ of $X_m$ is defined similarly by

$$\Omega_\alpha^\beta = d\omega_\alpha^\beta + \omega_\lambda^\beta \wedge \omega_\alpha^\lambda. \tag{7.28}$$

As in (6.4) the latter can be represented explicitly in the form

$$\Omega_\alpha^\beta = -\tfrac{1}{2} K_\alpha{}^\beta{}_{\varepsilon\gamma} \, du^\varepsilon \wedge du^\gamma, \tag{7.29}$$

in which $K_\alpha{}^\beta{}_{\varepsilon\gamma}$ denotes the curvature tensor of $X_m$ with respect to the given connection $\Gamma_\beta{}^\alpha{}_\gamma$ on $X_m$:

$$K_\alpha{}^\beta{}_{\varepsilon\gamma} = \frac{\partial \Gamma_\alpha{}^\beta{}_\varepsilon}{\partial u^\gamma} - \frac{\partial \Gamma_\alpha{}^\beta{}_\gamma}{\partial u^\varepsilon} + \Gamma_\lambda{}^\beta{}_\gamma \Gamma_\alpha{}^\lambda{}_\varepsilon - \Gamma_\lambda{}^\beta{}_\varepsilon \Gamma_\alpha{}^\lambda{}_\gamma. \tag{7.30}$$

The formula (7.27) holds for any $p$-form $\Pi_\alpha^j$ of the given type: in particular, therefore, it will hold for the 0-form $B_\alpha^j$ defined by (7.2), so that

$$D(DB_\alpha^j) = B_\alpha^h \Omega_h^j - B_\beta^j \Omega_\alpha^\beta. \tag{7.31}$$

Let us now define the quantities

$$H_\alpha{}^j{}_\beta = B_{\alpha\|\beta}^j, \tag{7.32}$$

which represent the components of a type $(1, 0)$ tensor relative to (7.3) and of a type $(0, 2)$ tensor relative to (7.4). Then

$$DB_\alpha^j = H_\alpha{}^j{}_\beta \, du^\beta, \tag{7.33}$$

and with the aid of (7.23) it is found that

$$\begin{aligned}
D(DB_\alpha^j) &= (DH_\alpha{}^j{}_\beta) \wedge du^\beta + H_\alpha{}^j{}_\beta D(du^\beta) \\
&= H_\alpha{}^j{}_{\beta\|\gamma} \, du^\gamma \wedge du^\beta - H_\alpha{}^j{}_\beta \Omega^\beta.
\end{aligned} \tag{7.34}$$

When this is substituted in (7.31), the latter assumes the form

$$H_\alpha{}^j{}_{\beta\|\gamma} \, du^\beta \wedge du^\gamma + H_\alpha{}^j{}_\beta \Omega^\beta = -B_\alpha^h \Omega_h^j + B_\beta^j \Omega_\alpha^\beta. \tag{7.35}$$

This is the identity which we have been seeking; it represents the relationship between the respective curvature 2-forms of $X_n$ and $X_m$ in terms of the second covariant derivatives of the quantities $B_\alpha^j$. In order to obtain a representation of (7.35) in terms of components, we substitute from (7.21), (7.29), and (6.4), which allows us to write (7.35) as

$$(H_\alpha{}^j{}_{\beta\|\gamma} - H_\alpha{}^j{}_{\gamma\|\beta}) \, du^\beta \wedge du^\gamma + H_\alpha{}^j{}_\varepsilon S_\beta{}^\varepsilon{}_\gamma \, du^\beta \wedge du^\gamma$$
$$= B_\alpha^h K_h{}^j{}_{lk} \, dx^l \wedge dx^k - B_\varepsilon^j K_\alpha{}^\varepsilon{}_{\beta\gamma} \, du^\beta \wedge du^\gamma,$$

or, if we apply (7.7) to the first term on the right-hand side,

$$H_\alpha{}^j{}_{\beta\|\gamma} - H_\alpha{}^j{}_{\gamma\|\beta} = K_l{}^j{}_{hk} B_\alpha^l B_\beta^h B_\gamma^k - B_\lambda^j K_\alpha{}^\lambda{}_{\beta\gamma} - H_\alpha{}^j{}_\varepsilon S_\beta{}^\varepsilon{}_\gamma. \tag{7.36}$$

This formula clearly exhibits the relationship between the respective curvature tensors of $X_n$ and $X_m$ in terms of the third derivatives of the embedding equations (7.1). For most practical purposes, the relation (7.36) is too general, since the curvature tensors are defined with respect to arbitrary connections on $X_n$ and $X_m$. However, it will be seen that, subject to certain restrictions to be discussed in Chapter 7, the formula (7.36) implies the famous equation of Gauss and Codazzi of classical differential geometry.

## PROBLEMS

5.1  (a) If $\omega$ is the 2-form given by

$$\omega = dx^1 \wedge dx^2 + dx^3 \wedge dx^4,$$

calculate $\omega \wedge \omega$, and $\omega \wedge \omega \wedge \omega$.
(b) If $\omega$ is the 2-form given by

$$\omega = dx^1 \wedge dx^2 + dx^3 \wedge dx^4 + \cdots + dx^{2m-1} \wedge dx^{2m}$$

calculate $\omega \wedge \omega \wedge \cdots \wedge \omega$ ($m$ times), and $\omega \wedge \omega \wedge \cdots \wedge \omega$ ($m + 1$ times).

5.2  If

$$f = (x^1)^2 + (x^2)^2 + (x^3)^2 \quad g = x^1 + x^2 + x^3$$

evaluate $df \wedge dg$.

5.3  If $\omega = x^1 x^3 \, dx^2 + x^1 x^2 \, dx^3 + (x^1 + x^3 x^5) \, dx^4 + x^3 x^4 \, dx^5$ evaluate $d\omega$, $\omega \wedge d\omega$, $d\omega \wedge d\omega$, and $\omega \wedge d\omega \wedge d\omega$.

5.4  If, in $E_3$,

$$\alpha = A_i \, dx^i$$
$$\beta = B_3 \, dx^1 \wedge dx^2 + B_1 \, dx^2 \wedge dx^3 + B_2 \, dx^3 \wedge dx^1,$$

show that

$$d\alpha = 0 \text{ if and only if curl } \mathbf{A} = 0,$$
$$d\beta = 0 \text{ if and only if div } \mathbf{B} = 0.$$

What are the identities $d(d\phi) = 0$ in vector notation, if $\phi$ is a 0-form?

5.5  In $E_4$, with $x^1 = x$, $x^2 = y$, $x^3 = z$, and $x^4 = ct$, let

$$\alpha = \sum_{\beta=1}^{3} E_\beta \, dx^\beta \wedge dx^4 + H_1 \, dx^2 \wedge dx^3 + H_3 \, dx^1 \wedge dx^2 + H_2 \, dx^3 \wedge dx^1$$

and

$$\omega = -\sum_{\beta=1}^{3} H_\beta \, dx^\beta \wedge dx^4 + E_1 \, dx^2 \wedge dx^3 + E_3 \, dx^1 \wedge dx^2 + E_2 \, dx^3 \wedge dx^1.$$

Show that the source-free Maxwell's equation given in Problem 4.5 with $\rho = 0$, $\mathbf{J} = \mathbf{0}$ can be expressed in the form

$$d\alpha = 0, \qquad d\omega = 0.$$

If

$$\lambda = \sum_{\beta=1}^{3} A_\beta \, dx^\beta - \phi \, dx^4$$

show that $d\lambda = \alpha$ if and only if

$$\mathbf{H} = \text{curl } \mathbf{A} \quad \text{and} \quad \mathbf{E} = -\text{grad } \phi - \frac{\partial \mathbf{A}}{\partial t}.$$

5.6  If

$$\omega = x \, dy \wedge dz - 2zf(y) \, dx \wedge dy + yf(y) \, dz \wedge dx$$

determine $f$ so that
(a) $d\omega = dx \wedge dy \wedge dz$,
(b) $d\omega = 0$.

5.7  If $u = u(x, y, z)$, $v = v(x, y, z)$, and

$$\omega = (y - vu_x) \, dx - (x + vu_y) \, dy + (1 - vu_z) \, dz,$$

where $u_x = \partial u/\partial x$ etc., show that

$$d\omega = 0$$

if and only if $u_z = v_z = 0$ and $v_y u_x - v_x u_y = 2$. If these conditions are satisfied prove that $du$, $dv$, and $\omega$ are linearly independent.

5.8  If $\omega = (y \cos xy) \, dx + (1 + x \cos xy) \, dy$ show that

$$d\omega = 0$$

and find $\phi(x, y)$ for which $\omega = d\phi$.

5.9  Show that if $P = P(x, y)$ and $Q = Q(x, y)$ then

$$P \, dx + Q \, dy = 0$$

is always integrable.

5.10  If $P = P(x, y, u, v)$ and $Q = Q(x, y, u, v)$ under what conditions is the system of differential equations

$$dx = P \, du - Q \, dv$$
$$dy = Q \, du + P \, dv$$

completely integrable?

5.11  Show that the system of differential equations

$$2xy \, dx - (y + z) \, dy = 0$$
$$xz \, dy - yz \, dz = 0$$

is completely integrable.

5.12  If $f = f(x^1, x^2, y, y_1, y_2)$, show that the two statements

$$\frac{\partial}{\partial x^i} \left( \frac{\partial f}{\partial y_i} \right) - \frac{\partial f}{\partial y} = 0$$

and

$$\frac{\partial f}{\partial y} \, dx^1 \wedge dx^2 = d\left[ \frac{\partial f}{\partial y_1} \, dx^2 - \frac{\partial f}{\partial y_2} \, dx^1 \right]$$

are equivalent.

5.13  In $X_n$ let $\omega^1, \ldots, \omega^r$ be $r$-linearly independent 1-forms ($r \leq n$). The $r$ 1-forms $\pi_1, \ldots, \pi_r$ are also linearly independent and satisfy $\omega^1 \wedge \pi_1 + \omega^2 \wedge \pi_2 + \cdots + \omega^r \wedge \pi_r = 0$. Show that

$$\pi_h = \sum_{i=1}^{r} a_{hi} \omega^i$$

where $a_{hi} = a_{ih}$ (Cartan's lemma).

*5.14  A $p$-form $\Omega$ is said to be divisible by a 1-form $\omega$ if there exists a $(p-1)$-form $\pi$ for which $\Omega = \pi \wedge \omega$. Show that $\Omega$ is divisible by $\omega$ if and only if $\Omega \wedge \omega = 0$.

5.15  In $X_n$ let $F_{ij}$ be the components of a skew-symmetric tensor of type $(0, 2)$ and let $A_i$, $B_j$ be the components of type $(0, 1)$ tensors. Show that

$$F_{ij} = A_i B_j - A_j B_i$$

if and only if

$$F_{ij} F_{hk} + F_{ik} F_{jh} + F_{ih} F_{kj} = 0.$$

Show that if $n = 2$ or 3 the latter condition is an identity whereas if $n = 4$ it is equivalent to $\det |F_{ij}| = 0$.

\*5.16 Let $a_{ij} = a_{ji}$ and $a_{ij} = a_{ij}(x^h)$. Prove that a necessary and sufficient condition for the existence of functions $u_i$ $(i = 1, \ldots, n)$ for which

$$a_{ij} = \frac{\partial u_i}{\partial x^j} + \frac{\partial u_j}{\partial x^i}$$

is

$$\left( \frac{\partial^2 a_{ij}}{\partial x^k \partial x^l} - \frac{\partial^2 a_{lj}}{\partial x^k \partial x^i} \right) dx^j \wedge dx^k = 0.$$

(*Hint*: use the converse of Poincaré's lemma twice.) These are the integrability conditions for (1.1.6).

5.17 If $\Phi$ is a $p$-form, $\Phi = a_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$, we define its Lie derivative with respect to $v^i$ by

$$£_v \Phi = (£_v a_{i_1 \cdots i_p}) dx^{i_1} \wedge \cdots \wedge dx^{i_p}.$$

Prove that

(a) $£_v(c_1 \Phi + c_2 \Omega) = c_1 £_v \Phi + c_2 £_v \Omega$,

(b) $£_v(\Phi \wedge \omega) = (£_v \Phi) \wedge \omega + \Phi \wedge £_v \omega$,

(c) $£_v(d\Phi) = d(£_v \Phi)$,

where $\Phi, \Omega$ are $p$-forms, $\omega$ is a $q$-form, and $c_1, c_2$ are constants.

5.18 Let $\Phi$ be a $p$-form, $\Phi = a_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$, and let $A^i$ be the components of a type $(1, 0)$ tensor. Define the $(p-1)$-form

$$\underline{A} \rfloor \Phi = \frac{1}{(p-1)!} \delta^{j_1 \cdots j_p}_{i_1 \cdots i_{p-1} h} a_{j_1 \cdots j_p} A^h dx^{i_1} \wedge \cdots \wedge dx^{i_{p-1}} \quad \text{for} \quad p \geq 1$$

and

$$\underline{A} \rfloor \Phi = 0 \quad \text{for} \quad p = 0.$$

Show that

(a) $\underline{A} \rfloor (\underline{A} \rfloor \Phi) = 0$;

(b) $\underline{A} \rfloor (\Phi \wedge \Psi) = (\underline{A} \rfloor \Phi) \wedge \Psi + (-1)^p \Phi \wedge \underline{A} \rfloor \Psi$ where $\Psi$ is a $q$-form;

(c) $\underline{A} \rfloor (£_A \Phi) = £_A(\underline{A} \rfloor \Phi)$;

(d) $d(\underline{A} \rfloor \Phi) + \underline{A} \rfloor (d\Phi) = £_A \Phi$.

5.19 In the notation of Problem 5.18 define

$$\tilde{\Phi} = a_{i_1 \cdots i_p}(dx^{i_1} - A^{i_1} dt) \wedge \cdots \wedge (dx^{i_p} - A^{i_p} dt).$$

Show that $\tilde{\Phi} = \Phi - (\underline{A} \rfloor \Phi) \wedge dt$.

5.20 If $A^h = A^h(x^j)$ are the components of a type $(1, 0)$ vector field in $X_n$ show that the partial differential equation

$$A^h \frac{\partial \phi}{\partial x^h} = 0$$

has $(n-1)$ independent solutions $\phi^1, \ldots, \phi^{n-1}$. By considering the change of coordinates $\bar{x}^i = \phi^i$, $i = 1, \ldots, n-1$, $\bar{x}^n = f$ where $f$ is any function of $x^i$ for which $|\partial \bar{x}^i / \partial x^j| \neq 0$ prove that there always exists a coordinate system in which all the components of $A^h$ except one are zero.

\*5.21 Show that condition (4.51) is equivalent to each of the following
(a) There exist 1-forms $\theta^i_j$ for which

$$d\omega^i = \theta^i_j \wedge \omega^j;$$

(b) There exists a 1-form $\lambda$ for which

$$d(\omega^1 \wedge \cdots \wedge \omega^n) = \lambda \wedge \omega^1 \wedge \cdots \wedge \omega^n.$$

\*5.22 Prove that (4.7) is completely integrable if and only if there exist functions $A^i_j$, $v^h$ for which

$$\omega^i = A^i_h \, dv^h.$$

[*Hint*: substitute (4.13) in (4.6) and note (4.11), (4.16), (4.17), and (4.22).]

5.23 If $\omega = A_i \, dx^i$ show that there exist $\alpha, \phi$ for which

$$A_i = \alpha \frac{\partial \phi}{\partial x^i}$$

if and only if $d\omega \wedge \omega = 0$.

5.24 If $\omega^1 = A_i \, dx^i$, $\omega^2 = B_j \, dx^j$ show that there exist $\alpha, \beta, \gamma, \delta, \phi$ and $\psi$ for which

$$A_i = \alpha \frac{\partial \phi}{\partial x^i} + \beta \frac{\partial \psi}{\partial x^i},$$

$$B_i = \gamma \frac{\partial \phi}{\partial x^i} + \delta \frac{\partial \psi}{\partial x^i},$$

if and only if $d\omega^\alpha \wedge \omega^1 \wedge \omega^2 = 0$ ($\alpha = 1, 2$). Express this latter condition in terms of $A_i$ and $B_j$.

5.25 Prove that

$$y \frac{\partial \phi}{\partial x} - x \frac{\partial \phi}{\partial y} = 0$$

$$x \frac{\partial \phi}{\partial x} - y \frac{\partial \phi}{\partial y} + \sin z \frac{\partial \phi}{\partial z} = 0$$

is completely integrable and has two functionally independent solutions.

5.26 In a particular $X_4$, $\omega_h{}^j$, defined by (2.20), are given by

$$\omega_h{}^j = \tfrac{1}{2}(\lambda \delta^1_h \delta^j_1 + \dot{\mu} \delta^4_h \delta^j_4) \, dx^1 + \tfrac{1}{2} \mu (\delta^1_h \delta^j_4 - e^{\mu - \lambda} \delta^4_h \delta^j_1) \, dx^4$$

where $\lambda, \mu$ are functions of $x^1$ alone and $\dot{\lambda} = d\lambda/dx^1$. Compute $\Omega_h{}^j$ and $\Omega^j$, defined by (6.3) and (6.5), respectively. Hence show that $K_i{}^j{}_{hk} = 0$ if and only if

$$2 \ddot{\mu} - \dot{\lambda} \dot{\mu} + \dot{\mu}^2 = 0.$$

5.27 If $\Omega = A_{ijk} \, dx^i \wedge dx^j \wedge dx^k$ show that there exist 1-forms $\omega^1$, $\omega^2$, $\omega^3$ for which $\Omega = \omega^1 \wedge \omega^2 \wedge \omega^3$ if and only if

$$A_{iji_1} A_{i_2 i_3 i_4} - A_{iji_2} A_{i_1 i_3 i_4} + A_{iji_3} A_{i_1 i_2 i_4} - A_{iji_4} A_{i_1 i_2 i_3} = 0,$$

assuming without loss of generality that $A_{ijk}$ are completely skew-symmetric. Prove that this is an identity if $n = 4$.

*5.28 If $\Omega$ is a 2-form and $\omega^1$ and $\omega^2$ are 1-forms show that

$$\Omega \wedge \omega^1 \wedge \omega^2 = 0$$

if and only if there exist 1-forms $\pi_1$ and $\pi_2$ for which

$$\Omega = \pi_1 \wedge \omega^1 + \pi_2 \wedge \omega^2.$$

# 6
# INVARIANT PROBLEMS
# IN THE CALCULUS
# OF VARIATIONS

A most fascinating phenomenon is the impact of the concept of invariance under coordinate transformations on the calculus of variations. This is due to the fact that when one assumes that the action integral of a variational principle is an invariant (as is usually done in physical field theories), it is necessary to stipulate that the integrand be a scalar density. The profound implications of this state of affairs were clearly recognized by Hilbert [1], and it is the objective of this chapter to describe these matters in some detail. Since it is not assumed that the reader is familiar with the calculus of variations, some of the elementary aspects of the latter are developed ab initio with the aid of tensor methods and the calculus of forms. Moreover, with the exception of the first part of Section 6.2, the fairly recent approach of Carathéodory to the calculus of variations is introduced in preference to the more widely known classical techniques based on the theory of the first and second variations. In this sense this chapter is to be regarded as an end in itself. However, it should be pointed out that for the purposes of Riemannian geometry as presented in Chapter 7, only an acquaintance with the contents of Sections 6.1 and 6.2 is required, whereas the material of Chapter 8 is dependent on Sections 6.1, 6.2, 6.5, and 6.7.

Particular emphasis is placed on the theorems of Noether, which are concerned with the implications of the assumption that the fundamental integral be invariant under a special $r$-parameter transformation group. These theorems are of considerable importance in field-theoretic applications since they establish the existence and precise nature of certain conservation

laws which result from the given invariance requirement. Instead of following the original derivation of Noether [1], which is fairly complicated and depends on some deep and difficult theorems in the calculus of variations, the approach presented here is simple and direct, being essentially an exercise in the techniques of the tensor calculus.

## 6.1  THE SIMPLEST PROBLEM IN THE CALCULUS OF VARIATIONS; INVARIANCE REQUIREMENTS

We begin with a brief formulation of the simplest problem in the calculus of variations† on our differentiable manifold $X_n$, after which the relevant invariance requirements will be considered in some detail.

A curve $C$ of $X_n$ is represented parametrically in the form

$$x^j = x^j(t), \tag{1.1}$$

in which $t$ denotes a given parameter. For the sake of simplicity we consider solely curves for which the representation (1.1) is given by class $C^2$ functions, and, as before, we use the notation

$$\dot{x}^j = \frac{dx^j}{dt} = \dot{x}^j(t). \tag{1.2}$$

Let us suppose that we are given a class $C^2$ function $L(t, x^j, \dot{x}^j)$ of $2n + 1$ independent variables which is to be referred to henceforth as the *Lagrangian* or *Lagrange function*. Furthermore, let $P_1$, $P_2$ be two points of $X_n$ which are joined by a curve such as (1.1), these two points corresponding to the parameter values $t_1$ and $t_2$, respectively. One can then construct the integral of the Lagrangian $L$ along the curve $C$, namely,

$$I(C) = \int_{t_1}^{t_2} L(t, x^j(t), \dot{x}^j(t)) \, dt, \tag{1.3}$$

where it is to be understood that the functions $x^j(t)$, $\dot{x}^j(t)$ which appear as arguments in the integrand are obtained from the relations (1.1) and (1.2) which refer to the curve $C$. Clearly the value of the integrand and hence also that of the integral will in general depend on the functions $x^j(t)$, $\dot{x}^j(t)$; in other words, the integral (1.3) generally depends on the choice of the curve $C$ joining the points $P_1$ and $P_2$. The resulting problem in the calculus of variations is concerned with the conditions which a curve $C$ must satisfy in order that it afford an extreme value to the integral (1.3) as compared with all other curves possessing the same end-points $P_1$ and $P_2$.

† More extensive treatments of the calculus of variations are presented, for instance, in the following texts: Akhiezer [1], Carathéodory [1], Gelfand and Fomin [1], and Rund [2].

Variational problems of this kind occur in many branches of pure and applied mathematics. Two simple examples are presented here.

EXAMPLE 1

In an $n$-dimensional Euclidean space $E_n$ referred to rectangular coordinates the length of an arc of a curve $C$ between two points $P_1$ and $P_2$ is given by

$$s = \int_{t_1}^{t_2} \left( \sum_{j=1}^{n} dx^j \, dx^j \right)^{1/2} = \int_{t_1}^{t_2} \left( \sum_{j=1}^{n} \dot{x}^j \dot{x}^j \right)^{1/2} dt. \tag{1.4}$$

Intuitively it is obvious that the curve which affords a minimum to the integral (1.4) is the (unique) straight-line segment joining $P_1$, $P_2$.

EXAMPLE 2

Let $x^j$ denote the generalized coordinates of a classical dynamical system with $n$ degrees of freedom; the derivatives $\dot{x}^j = dx^j/dt$ may then be regarded as the generalized velocity components, provided that the parameter $t$ represents the time. The kinetic energy $T$ of such a system can usually be expressed as a quadratic form

$$T = \tfrac{1}{2} a_{hj} \dot{x}^h \dot{x}^j, \tag{1.5}$$

whose coefficients are given functions of the variables $x^h$. If it is assumed that the dynamical system is conservative, one may associate with it a potential function $V(x^h)$. A suitable Lagrangian may then be defined by putting

$$L = T - V, \tag{1.6}$$

which gives rise to a problem in the calculus of variations as exemplified by the integral (1.3). According to *Hamilton's Principle*, the motion of the dynamical system is described by curves in the configuration space $X_n$ which are such as to afford extreme values to this integral. (It should be emphasized that this is a very superficial formulation of Hamilton's Principle; see, e.g., Goldstein [1].) On the basis of this assumption the laws of motion governing our dynamical system may be derived by means of the techniques of the calculus of variations to be discussed below.

We now turn to the transformation properties of the fundamental integral (1.3) of our variational problem. Two totally distinct types of transformations must be considered:

1.  A coordinate transformation of the type

$$\bar{x}^j = \bar{x}^j(x^h), \tag{1.7}$$

which does not affect the parameter $t$.

2.   A parameter transformation of the type

$$\tau = \tau(t), \tag{1.8}$$

which does not affect the coordinates $x^h$.

It is assumed henceforth that the integral (1.3) is invariant under (1.7); we shall subsequently investigate the implications of the additional requirement that (1.3) be invariant also under (1.8).

As before, it is supposed that the transformation (1.7) possesses an inverse

$$x^h = x^h(\bar{x}^j). \tag{1.9}$$

The derivatives (1.2) constitute the components of a contravariant vector:

$$\dot{x}^h = \frac{\partial x^h}{\partial \bar{x}^j} \dot{\bar{x}}^j = \dot{x}^h(\bar{x}^j, \dot{\bar{x}}^j), \tag{1.10}$$

as is immediately evident by differentiation of (1.9). Furthermore, it follows directly from (1.10) that

$$\frac{\partial \dot{x}^h}{\partial \dot{\bar{x}}^j} = \frac{\partial x^h}{\partial \bar{x}^j}, \tag{1.11}$$

together with

$$\frac{\partial \dot{x}^h}{\partial \bar{x}^j} = \frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l} \dot{\bar{x}}^l. \tag{1.12}$$

Clearly the assumption that the integral (1.3) be invariant under (1.7) is equivalent to the stipulation that the Lagrangian $L$ be a scalar relative to (1.7). Thus, if we denote the transform of $L$ under (1.7) by $\bar{L}$, our invariance requirement is equivalent to the relation

$$\bar{L}(t, \bar{x}^j, \dot{\bar{x}}^j) = L(t, x^h(\bar{x}^j), \dot{x}^h(\bar{x}^j, \dot{\bar{x}}^j)), \tag{1.13}$$

where, on the right-hand side, we have substituted from (1.9) and (1.10). This must be valid for all values of $(\bar{x}^j, \dot{\bar{x}}^j)$. Thus, differentiating with respect to $\dot{\bar{x}}^j$, and using (1.11), we obtain

$$\frac{\partial \bar{L}}{\partial \dot{\bar{x}}^j} = \frac{\partial L}{\partial \dot{x}^h} \frac{\partial \dot{x}^h}{\partial \dot{\bar{x}}^j} = \frac{\partial L}{\partial \dot{x}^h} \frac{\partial x^h}{\partial \bar{x}^j}, \tag{1.14}$$

from which it is evident that *the derivatives $\partial L/\partial \dot{x}^h$ constitute the components of a covariant vector*. However, if we differentiate (1.13) with respect to $\bar{x}^j$ and subsequently apply (1.12), we find that

$$\frac{\partial \bar{L}}{\partial \bar{x}^j} = \frac{\partial L}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^j} + \frac{\partial L}{\partial \dot{x}^h} \frac{\partial \dot{x}^h}{\partial \bar{x}^j} = \frac{\partial L}{\partial x^h} \frac{\partial x^h}{\partial \bar{x}^j} + \frac{\partial L}{\partial \dot{x}^h} \frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l} \dot{\bar{x}}^l, \tag{1.15}$$

which indicates that *the derivatives $\partial L/\partial x^h$ do not form the components of a tensor*. In this connection it should be recalled that the derivatives $\partial \phi/\partial x^h$ of a scalar function $\phi(x^h)$ represent a covariant vector, namely, the gradient of $\phi$: however, in the case of the scalar Lagrangian $L(t, x^h, \dot{x}^h)$ this is not true as a result of the dependence of $L$ on the variables $\dot{x}^h$.

It is therefore desirable to construct a covariant vector associated with the derivatives $\partial L/\partial x^h$ which would represent a generalized gradient of $L$. This is easily done by eliminating the second term on the right-hand side of (1.15). To this end we differentiate (1.14) with respect to $t$, which gives

$$\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{\bar{x}}^j}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^h}\right)\frac{\partial x^h}{\partial \bar{x}^j} + \frac{\partial L}{\partial \dot{x}^h}\frac{d}{dt}\left(\frac{\partial x^h}{\partial \bar{x}^j}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^h}\right)\frac{\partial x^h}{\partial \bar{x}^j} + \frac{\partial L}{\partial \dot{x}^h}\frac{\partial^2 x^h}{\partial \bar{x}^j \partial \bar{x}^l}\dot{\bar{x}}^l. \tag{1.16}$$

When (1.15) is subtracted from this result it is found that

$$\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{\bar{x}}^j}\right) - \frac{\partial \bar{L}}{\partial \bar{x}^j} = \frac{\partial x^h}{\partial \bar{x}^j}\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^h}\right) - \frac{\partial L}{\partial x^h}\right], \tag{1.17}$$

from which it is evident that *the quantities defined by*

$$E_j(L) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) - \frac{\partial L}{\partial x^j} \tag{1.18}$$

*constitute the components of a covariant vector*. Clearly this vector may be interpreted as a generalized gradient of the Lagrangian $L$: it will be seen that it plays a fundamental role in all subsequent developments. We shall henceforth refer to $E_j(L)$ as the *Euler–Lagrange vector of $L$*; written out in full it is seen to possess the form

$$E_j(L) = \frac{\partial^2 L}{\partial \dot{x}^j \partial \dot{x}^h}\ddot{x}^h + \frac{\partial^2 L}{\partial \dot{x}^j \partial x^h}\dot{x}^h + \frac{\partial^2 L}{\partial \dot{x}^j \partial t} - \frac{\partial L}{\partial x^j}. \tag{1.19}$$

It should be noted that, in contrast to the vector (1.14), the Euler–Lagrange vector depends on the second derivatives $\ddot{x}^j$ of the functions (1.1).

Let us now turn to the parameter transformation (1.8), of which it will be assumed that the functions $\tau(t)$ are of class $C^2$ and such that

$$\dot{\tau} = \frac{d\tau}{dt} > 0. \tag{1.20}$$

The integral (1.3) is said to be *parameter-invariant*, if, for *any* curve $C$ on $X_n$ joining *any* points $P_1$ and $P_2$ we have

$$\int_{C \, P_1}^{P_2} L\left(\tau, x^j, \frac{dx^j}{d\tau}\right) d\tau = \int_{C \, P_1}^{P_2} L\left(t, x^j, \frac{dx^j}{dt}\right) dt. \tag{1.21}$$

Clearly this condition can be satisfied if and only if the Lagrangian $L$ is such that

$$L\left(\tau, x^j, \frac{dx^j}{d\tau}\right) d\tau = L\left(t, x^j, \frac{dx^j}{dt}\right) dt,$$

or

$$L\left(\tau, x^j, \frac{\dot{x}^j}{\dot{t}}\right)\dot{t} = L(t, x^j, \dot{x}^j), \tag{1.22}$$

for arbitrary values of the arguments $(x^j, \dot{x}^j)$, and for any function $\tau = \tau(t)$. In particular, if the latter is specified by $\tau = t + \alpha$ for some arbitrary constant $\alpha$, we have $\dot{t} = 1$, and (1.22) reduces to

$$L(t + \alpha, x^j, \dot{x}^j) = L(t, x^j, \dot{x}^j).$$

In this case the left-hand side apparently depends on $\alpha$, while this is not the case for the right-hand side, which implies that we must have

$$\frac{\partial L}{\partial t} = 0, \tag{1.23}$$

that is, the Lagrangian $L$ may not depend explicitly on $t$. Assuming this to be the case generally, and putting $\lambda = \dot{t}^{-1}$, it is inferred that (1.22) is equivalent to

$$L(x^j, \lambda\dot{x}^j) = \lambda L(x^j, \dot{x}^j), \tag{1.24}$$

in which $\lambda$ is essentially arbitrary except for the restriction $\lambda > 0$, which is a direct consequence of (1.20). Accordingly the Lagrangian $L(x^j, \dot{x}^j)$ must be positively homogeneous of the first degree in the $\dot{x}^j$.

The requirements (1.23) and (1.24) thus emerge as necessary conditions for the parameter-invariance of the fundamental integral (1.3). The sufficiency of these conditions is immediately evident from the following:

$$\int_C^{P_2}{}_{P_1} L\left(x^j, \frac{dx^j}{d\tau}\right) d\tau = \int_C^{P_2}{}_{P_1} L\left(x^j, \frac{dx^j}{d\tau}\dot{t}\right) \frac{d\tau}{\dot{t}} = \int_C^{P_2}{}_{P_1} L\left(x^j, \frac{dx^j}{dt}\right) dt,$$

where, in the first step, we have used (1.24).

*Remark 1.*  In the calculus of variations it is usually assumed that the fundamental integral (1.3) is invariant under the coordinate transformation (1.7). However, in many applications the parameter-invariance of (1.3) cannot be presupposed. For instance, the Lagrangian defined by (1.6) and (1.5) does *not* satisfy the homogeneity condition (1.24). This phenomenon is, in fact, quite natural, since in the example (1.6) the parameter $t$ refers to the

time in the sense of classical mechanics, in which the existence of an absolute time scale is presupposed, which obviates the need for transformations of the type (1.8). (This is in direct contrast to the requirements of relativistic mechanics.)

*Remark 2.*  In the case of most differential-geometric applications one is confronted with parameter-invariant integrals. For instance, the integral (1.4) is of this type, and it will be seen that this is true of many examples encountered below.

*Remark 3.*  Let us suppose, for the moment, that we are concerned with a parameter-invariant integral, so that the validity of (1.23) and (1.24) may be assumed. By means of Euler's theorem on homogeneous functions we then infer that

$$\frac{\partial L(x^h, \dot{x}^h)}{\partial \dot{x}^j} \dot{x}^j = L(x^h, \dot{x}^h). \tag{1.25}$$

However, it then follows from (1.18) that

$$E_j(L)\dot{x}^j = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j} \dot{x}^j\right) - \frac{\partial L}{\partial \dot{x}^j} \ddot{x}^j - \frac{\partial L}{\partial x^j} \dot{x}^j = \frac{dL}{dt} - \frac{dL}{dt} = 0 \tag{1.26}$$

identically. Since $E_j(L)$ is a covariant vector, we interpret this result by saying that *the Euler-Lagrange vector along any curve is normal to the tangent vector of that curve.* For instance, in the case of the integral (1.4) we have $L = (\sum_{j=1}^n \dot{x}^j\dot{x}^j)^{1/2}$, so that

$$E_j(L) = \frac{\ddot{x}^j}{(\sum_{h=1}^n \dot{x}^h\dot{x}^h)^{1/2}} - \frac{\dot{x}^j(\sum_{k=1}^n \dot{x}^k\ddot{x}^k)}{(\sum_{h=1}^n \dot{x}^h\dot{x}^h)^{3/2}},$$

which coincides in direction with the principal normal of the curve under consideration.

## 6.2  FIELDS OF EXTREMALS

In order that a curve $C$ of $X_n$ afford an extreme value to the fundamental integral (1.3) it is necessary that the functions $x^j(t)$ which define $C$ satisfy a system of $n$ second-order ordinary differential equations, the so-called Euler–Lagrange equations. In this section these equations are derived by means of two entirely distinct techniques, of which the first is the traditional one, while the second is deeper and more elegant and therefore conceptually somewhat more difficult.

### First Method

Let $C: x^j = x^j(t)$ be a given class $C^1$ curve of $X_n$ joining the points $P_1$ and $P_2$, which correspond to the parameter values $t_1$ and $t_2$, respectively. Let us construct a one-parameter family of curves $C(u)$ by writing

$$x^j(t, u) = x^j(t) + u\zeta^j(t), \tag{2.1}$$

in which $u$ denotes the parameter, while the $\zeta^j(t)$ are class $C^1$ functions of $t$ such that

$$\zeta^j(t_1) = \zeta^j(t_2) = 0. \tag{2.2}$$

The latter conditions ensure that each curve of the family passes through the points $P_1$ and $P_2$. Clearly the given curve $C$ is the member of the family which corresponds to the parameter value $u = 0$. Furthermore, differentiation of (2.1) with respect to $t$ yields

$$\dot{x}^j(t, u) = \dot{x}^j(t) + u\dot{\zeta}^j(t), \tag{2.3}$$

where $\dot{x}^j(t, u)$ denotes the partial derivatives $\partial x^j(t, u)/\partial t$.

The value of the fundamental integral (1.3) evaluated along the curve $C(u)$ between the points $P_1$ and $P_2$ is denoted by $I(u)$:

$$I(u) = \int_{t_1}^{t_2} L\{t, x^j(t) + u\zeta^j(t), \dot{x}^j(t) + u\dot{\zeta}^j(t)\} \, dt. \tag{2.4}$$

It is generally possible to express $I(u)$ as a power series in $u$ as follows:

$$I(u) = I(0) + \delta I + \delta^2 I + \cdots, \tag{2.5}$$

where we have used the following notation which is customary in the calculus of variations:

$$\delta I = u\left(\frac{dI}{du}\right)_{u=0}, \tag{2.6}$$

$$\delta^2 I = \tfrac{1}{2}u^2\left(\frac{d^2I}{du^2}\right)_{u=0}, \tag{2.7}$$

these quantities usually being referred to as the *first* and *second variations* of the fundamental integral (1.3).

Now let us suppose that the curve $C$ corresponding to the parameter value $u = 0$ affords an extreme value to our integral; that is, it is assumed that $I(u)$ takes on an extreme value for $u = 0$. This implies that the condition

$$\delta I = 0 \tag{2.8}$$

must necessarily be satisfied. In order to exploit this requirement we obviously have to evaluate the first variation (2.6).

From (2.1) and (2.3) we have

$$\frac{\partial x^j}{\partial u} = \zeta^j, \qquad \frac{\partial \dot{x}^j}{\partial u} = \dot{\zeta}^j,$$

and hence differentiation of (2.4) yields

$$\left(\frac{dI}{du}\right)_{u=0} = \int_{t_1}^{t_2} \left\{\frac{\partial L}{\partial x^j}\zeta^j + \frac{\partial L}{\partial \dot{x}^j}\dot{\zeta}^j\right\} dt. \tag{2.9}$$

The first term in the integrand on the right-hand side may be integrated by parts as follows:

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial x^j}\zeta^j \, dt = \left[\zeta^j \int_{t_1}^t \frac{\partial L}{\partial x^j} \, dt\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left\{\dot{\zeta}^j \int_{t_1}^t \frac{\partial L}{\partial x^j} \, dt\right\} dt. \tag{2.10}$$

However, because of (2.2) the first term on the right-hand side of this expression vanishes identically. Accordingly the substitution of (2.10) in (2.9) gives

$$\left(\frac{dI}{du}\right)_{u=0} = \int_{t_1}^{t_2} \left\{\frac{\partial L}{\partial \dot{x}^j} - \int_{t_1}^t \frac{\partial L}{\partial x^j} \, dt\right\}\dot{\zeta}^j \, dt, \tag{2.11}$$

and thus, because of (2.6), the condition (2.8) is equivalent to

$$\int_{t_1}^{t_2} \Phi_j \dot{\zeta}^j \, dt = 0, \tag{2.12}$$

where, for the sake of brevity, we have put

$$\Phi_j = \frac{\partial L}{\partial \dot{x}^j} - \int_{t_1}^t \frac{\partial L}{\partial x^j} \, dt. \tag{2.13}$$

A set of $n$ constants $c_j$ is now defined by the relations

$$c_j(t_2 - t_1) = \int_{t_1}^{t_2} \Phi_j \, dt, \tag{2.14}$$

so that

$$\int_{t_1}^{t_2} (\Phi_j - c_j) \, dt = 0. \tag{2.15}$$

We now choose the functions $\zeta^j(t)$ which appear in (2.1) as follows:

$$\zeta^j(t) = \int_{t_1}^t (\Phi_j - c_j) \, dt, \tag{2.16}$$

it being obvious from (2.16) that the conditions (2.2) are satisfied for these functions. Moreover,

$$\dot{\zeta}^j(t) = \Phi_j - c_j, \tag{2.17}$$

so that (2.12) is equivalent to

$$\int_{t_1}^{t_2} \sum_{j=1}^{n} \Phi_j(\Phi_j - c_j)\, dt = 0,$$

or, because of (2.15),

$$\int_{t_1}^{t_2} \sum_{j=1}^{n} (\Phi_j - c_j)(\Phi_j - c_j)\, dt = 0. \tag{2.18}$$

The integrand of this integral consists of a sum of squares. Thus (2.18) implies that each of these terms vanishes, that is, that

$$\Phi_j - c_j = 0,$$

or, in view of (2.13),

$$\frac{\partial L}{\partial \dot{x}^j} - \int_{t_1}^{t} \frac{\partial L}{\partial x^j}\, dt = c_j. \tag{2.19}$$

This relation is differentiated with respect to $t$, which yields

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^j}\right) - \frac{\partial L}{\partial x^j} = 0, \tag{2.20}$$

or, in terms of the notation (1.18)

$$E_j(L) = 0. \tag{2.21}$$

We have therefore established the following important result: *In order that the curve C afford an extreme value to the fundamental integral* (1.3) *it is necessary that the functions $x^j(t)$ defining C are such that the Euler–Lagrange vector $E_j(L)$ vanishes along C.* The equations (2.20) or (2.21) which express this condition are usually referred to as the *Euler–Lagrange equations*: it is evident from (1.19) that these consist of a system of $n$ second-order ordinary differential equations. Any curve satisfying the differential equations (2.21) is called an *extremal*.

*Remark 1.* Since $E_j(L)$ is a covariant vector, as was established in the previous section, it follows that the Euler–Lagrange equations are invariant under coordinate transformations.

*Remark 2.* The above derivation of the Euler–Lagrange equations is valid irrespective of the parameter-invariance or otherwise of the fundamental integral.

*Remark 3.* It should be stressed that the Euler–Lagrange equations are merely necessary conditions for an extreme value of the fundamental integral. There are many examples which illustrate the fact that the Euler–Lagrange equations do *not* represent sufficient conditions for extreme values.

### Second Method (Carathéodory [1])

We now consider an alternative approach to the Euler–Lagrange equations. In order to be able to present this approach in a manner which clearly demonstrates its conceptual significance, it is necessary to introduce some important new ideas. In particular, it is advisable to base our construction on the $(n + 1)$-dimensional manifold $X_{n+1}$, whose coordinates are represented by the variables $(t, x^j)$, where the $x^j$ refer as before to the given manifold $X_n$. A curve $C$ of $X_{n+1}$ can be represented again by equations of the type $x^j = x^j(t)$; however, it should be stressed that in this representation the parameter $t$ has been identified with the $(n + 1)$th coordinate of $X_{n+1}$. This is not a serious drawback, for, as will be seen almost immediately, the theory to be developed below is applicable solely to nonparameter-invariant integrals.

According to (1.14) we may always associate a covariant vector with a scalar Lagrangian $L$. In fact, we shall now write

$$p_j = \frac{\partial L(t, x^h, \dot{x}^h)}{\partial \dot{x}^j}, \tag{2.22}$$

these quantities being called the components of the *canonical momentum*. [This nomenclature arises from the fact that for the Lagrangian (1.6), which refers to a dynamical system, the quantities (2.22) represent the components of the generalized momentum of that system.] Furthermore, it will be assumed that one can solve the system (2.22) for the $n$ components $\dot{x}^j$ as functions of the $p_j$, that is, that it is possible to write

$$\dot{x}^j = \phi^j(t, x^h, p_h). \tag{2.23}$$

In fact, under these circumstances (2.22) represents a one-to-one correspondence between the covariant vectors $p_j$ and the contravariant vectors $\dot{x}^j$. It should be emphasized that this assumption is tantamount to the requirement that

$$\det\left(\frac{\partial^2 L}{\partial \dot{x}^j\, \partial \dot{x}^h}\right) \neq 0 \tag{2.24}$$

by virtue of the inverse function theorem.

This assumption excludes homogeneous Lagrangians which satisfy condition (1.25), for the latter implies that the determinant which appears in (2.24) vanishes identically, as is easily verified by differentiation of (1.25). Thus the present approach must be modified for parameter-invariant problems in the calculus of variations; however, this modification will not be discussed here.

With the aid of (2.23) one can now associate the so-called *Hamiltonian function* $H(t, x^h, p_h)$ with the Lagrangian $L(t, x^h, \dot{x}^h)$ as follows:

$$H(t, x^h, p_h) = -L[t, x^h, \phi^h(t, x^l, p_l)] + p_j \phi^j(t, x^h, p_h). \qquad (2.25)$$

Since $p_j \phi^j = p_j \dot{x}^j$ is an inner product it follows that the function thus defined is also a scalar.

Again, this definition is motivated by known results from classical analytical dynamics. The kinetic energy $T$ which appears in the special Lagrangian (1.6) is given by the quadratic form (1.5) which must necessarily be positive definite if $T$ is to be always positive. This implies that the determinant of the coefficients $a_{hk}$ of (1.5) be positive, which in turn guarantees the existence of an inverse $a^{hj}$:

$$a_{hk} a^{jh} = \delta_k^j.$$

Corresponding to (2.22) we have, for the special Lagrangian (1.6):

$$p_j = a_{hj} \dot{x}^h,$$

the counterpart of (2.23) being

$$\dot{x}^j = a^{jh} p_h.$$

The assumption (2.24) is automatically satisfied since in this case

$$\det\left(\frac{\partial^2 L}{\partial \dot{x}^j \partial \dot{x}^h}\right) = \det(a_{jh}) > 0.$$

Because of (1.6) and the preceding relations the definition (2.25) is equivalent to

$$H(t, x^h, p_h) = -\tfrac{1}{2} a_{hj} a^{hl} a^{jk} p_l p_k + V + a^{jh} p_h p_j = \tfrac{1}{2} a^{hj} p_h p_j + V,$$

so that in this case the Hamiltonian represents the total energy $T + V$ of the dynamical system described by the Lagrangian (1.6).

The Hamiltonian (2.25) satisfies certain important identities. Let us differentiate (2.25) with respect to $p_j$:

$$\frac{\partial H}{\partial p_j} = -\frac{\partial L}{\partial \dot{x}^h} \frac{\partial \phi^h}{\partial p_j} + p_h \frac{\partial \phi^h}{\partial p_j} + \phi^j.$$

As a result of (2.22) the first two terms on the right-hand side cancel; thus

$$\frac{\partial H}{\partial p_j} = \phi^j(t, x^h, p_h) = \dot{x}^j. \qquad (2.26)$$

Similarly it is easily seen that

$$\frac{\partial H}{\partial x^j} = -\frac{\partial L}{\partial x^j}, \qquad (2.27)$$

and

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \qquad (2.28)$$

Clearly (2.26) represents the components of a contravariant vector, while (2.27) is a tensor equation despite the fact that the individual derivatives $\partial H/\partial x^j$, $\partial L/\partial x^j$ are nontensorial. However, it is immediately evident from (1.18), (2.22), and (2.27) that the Euler–Lagrange vector $E_j(L)$ can now be expressed in the form

$$E_j(L) = \frac{dp_j}{dt} + \frac{\partial H}{\partial x^j}, \qquad (2.29)$$

which in turn implies the invariant nature of the identity (2.27).

Now let us suppose that we are given a scalar class $C^2$ function $S = S(t, x^h)$. Along a curve $C: x^j = x^j(t)$ of $X_{n+1}$ we can form the total derivative

$$\frac{dS}{dt} = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial x^j} \dot{x}^j. \qquad (2.30)$$

With the aid of this quantity we may construct an alternative Lagrangian by writing

$$L^*(t, x^h, \dot{x}^h) = L(t, x^h, \dot{x}^h) - \frac{\partial S}{\partial t} - \frac{\partial S}{\partial x^j} \dot{x}^j, \qquad (2.31)$$

and accordingly one obtains a new integral, namely,

$$I^*(C) = \int_{P_1}^{P_2} L^*(t, x^h, \dot{x}^h)\, dt, \qquad (2.32)$$

which is to be evaluated along $C$ between the points $P_1$ and $P_2$. However, it is immediately evident from (2.31) that

$$I^*(C) = I(C) - (S_2 - S_1), \qquad (2.33)$$

where $S_2, S_1$ denote the values of the function $S(t, x^j)$ at $P_2$ and $P_1$ respectively. Since the difference

$$I^*(C) - I(C) = S_1 - S_2$$

is independent of the choice of the curve $C$ joining the points $P_1$ and $P_2$, it follows that $C$ will afford an extreme value to the integral $I^*$ if and only if it affords an extreme value to the integral $I$. Accordingly the integrals $I^*$ and $I$ are called *equivalent integrals* in the calculus of variations. The so-called method of equivalent integrals, which is to be briefly sketched below, entails the construction of curves which afford extreme values to $I^*(C)$;

if this has been achieved, the same curves will obviously also represent the solutions of our original problem.

To this end we now introduce the concept of a *geodesic field*. The latter is defined to be a field of contravariant vectors $\dot{x}^j = \psi^j(t, x^h)$, given at each point of a finite region $G$ of $X_{n+1}$ by a set of $n$ class $C^2$ functions $\psi^j(t, x^h)$ which are such that, for a suitably chosen function $S = S(t, x^h)$, the following conditions are satisfied:

$$L^*(t, x^h, \dot{x}^h) = 0 \quad \text{whenever} \quad \dot{x}^j = \psi^j(t, x^h), \qquad (2.34)$$

while

$$L^*(t, x^h, \dot{x}^h) > 0 \quad \text{otherwise.} \qquad (2.35)$$

It is not immediately evident that such a geodesic field can always be constructed, and we shall presently discuss the conditions which must be satisfied in order that this be possible. For the moment, however, let us assume that we are in possession of a geodesic field, in which case the latter gives rise directly to a solution of our problem.

Regarded as a system of $n$ first-order ordinary differential equations, the relations

$$\dot{x}^j = \psi^j(t, x^h) \qquad (2.36)$$

may be integrated to yield a family of curves which cover the region $G$ of $X_{n+1}$ simply. Let $\Gamma: x^j = x^j(t)$ be a member of this family, and let the points $P_1, P_2$ on $\Gamma$ correspond to the parameter values $t_1$ and $t_2$ respectively. The tangent vector $\dot{x}^j$ of $\Gamma$ satisfies (2.36) at each point by construction. It then follows from (2.34) that

$$\int_{\Gamma\, t_1}^{t_2} L^*(t, x^h, \dot{x}^h)\, dt = 0. \qquad (2.37)$$

On the other hand, for any other curve $K$ in $G$ which joins the points $P_1$ and $P_2$ we have, by virtue of (2.35),

$$\int_{K\, t_1}^{t_2} L^*(t, x^h, \dot{x}^h)\, dt > 0. \qquad (2.38)$$

Thus the curve $\Gamma$ affords a minimum to the integral $I^*$, and hence, in view of our foregoing remarks, it affords a minimum to the given integral $I$. Accordingly the curve $\Gamma$ represents a solution of our problem in the calculus of variations. [This particular method entails the construction of a minimum; a maximum could have been obtained in the same manner by merely reversing the inequality in (2.35).]

Let us now investigate some of the basic properties of our geodesic field. From (2.34) it follows that $L^*(t, x^h, \dot{x}^h)$ assumes a minimum value for $\dot{x}^j = \psi^j(t, x^h)$. Thus $\partial L^*/\partial \dot{x}^j = 0$ for these values of $\dot{x}^j$. However, from (2.31) and (2.22) we have quite generally that

$$\frac{\partial L^*}{\partial \dot{x}^j} = \frac{\partial L}{\partial \dot{x}^j} - \frac{\partial S}{\partial x^j} = p_j - \frac{\partial S}{\partial x^j},$$

so that, for the geodesic field,

$$p_j = \frac{\partial S}{\partial x^j}. \qquad (2.39)$$

This defines the covariant vector field $p_j = p_j(t, x^h)$ as a function of position on the region $G$ of $X_{n+1}$, and by virtue of (2.23) the directional arguments

$$\dot{x}^j = \phi^j(t, x^h, p_h) = \psi^j(t, x^h) \qquad (2.40)$$

corresponding to the geodesic field are thus determined. These are substituted in the condition (2.34), after which (2.39) is applied, which gives

$$L(t, x^h, \phi^h) = \frac{\partial S}{\partial t} + \frac{\partial S}{\partial x^j} \phi^j = \frac{\partial S}{\partial t} + p_j \phi^j,$$

or, in terms of the definition (2.25),

$$\frac{\partial S}{\partial t} + H(t, x^h, p_h) = 0.$$

Again, because of (2.39), this can be written in the form

$$\frac{\partial S}{\partial t} + H\left(t, x^h, \frac{\partial S}{\partial x^h}\right) = 0. \qquad (2.41)$$

This is a first-order partial differential equation for our scalar function $S = S(t, x^j)$: clearly our construction of the geodesic field entails that $S$ be a solution of this equation, as is assumed henceforth. The partial differential equation (2.41) is of fundamental importance to all subsequent developments: it is known as the *Hamilton–Jacobi equation*.

The above argument shows that, if the function $S(t, x^h)$ is to define a geodesic field, it must necessarily satisfy the Hamilton–Jacobi equation. Conversely, for any given solution of this equation, the corresponding covariant vector field $p_j = \partial S/\partial x^j$ is determined as a function of position, and hence a unique contravariant vector field $\dot{x}^j = \psi^j(t, x^h)$ is obtained by means of (2.40). Because of (2.39) and (2.41) it is immediately evident that this vector field is such that the first condition on the geodesic field, namely (2.34), is satisfied. [The second condition, namely, (2.35), will be investigated presently.]

Any curve $\Gamma$ which represents a solution of the system (2.36), satisfies a system of first-order differential equations which is easily obtained as follows. Differentiating (2.39) along $\Gamma$ with respect to $t$, we obtain

$$\frac{dp_j}{dt} = \frac{\partial^2 S}{\partial t\, \partial x^j} + \frac{\partial^2 S}{\partial x^h\, \partial x^j} \dot{x}^h. \tag{2.42}$$

However, if the Hamilton–Jacobi equation (2.41) is differentiated partially with respect to $x^j$, it is seen that

$$\frac{\partial^2 S}{\partial t\, \partial x^j} + \frac{\partial H}{\partial x^j} + \frac{\partial H}{\partial p_h} \frac{\partial^2 S}{\partial x^h\, \partial x^j} = 0$$

or, if we invoke the identity (2.26):

$$\frac{\partial^2 S}{\partial t\, \partial x^j} + \frac{\partial^2 S}{\partial x^h\, \partial x^j} \dot{x}^h = -\frac{\partial H}{\partial x^j}.$$

When this is compared with (2.42), one obtains the so-called *canonical equations*

$$\frac{dp_j}{dt} = -\frac{\partial H}{\partial x^j}, \tag{2.43}$$

which, together with (2.26), represent *a system of $2n$ first-order ordinary differential equations which the curve $\Gamma$ must satisfy*. Moreover, from (2.29) it is evident that *the canonical equations* (2.43) *are equivalent to the Euler–Lagrange equations*. In fact, the latter result formally from (2.43) and (2.27) by elimination of the variables $p_j$. One therefore has the choice between a system of $2n$ first-order equations for the $2n$ variables $p_j$ and $x^j$, or a system of $n$ second-order equations for the $n$ variables $x^j$. For many theoretical purposes it is preferable to use the former system.

Finally, let us briefly glance at the conditions which must be satisfied in order that the validity of the inequality (2.35) is ensured. In terms of (2.31) this condition may be written in the form

$$L(t, x^h, \dot{x}^h) - \frac{\partial S}{\partial t} - \frac{\partial S}{\partial x^j} \dot{x}^j \geq 0,$$

with equality if and only if $\dot{x}^j = \psi^j(t, x^h)$. One may eliminate $\partial S/\partial t$ from this formulation by means of (2.41), where it must be remembered, however, that the latter applies to the geodesic field. This gives

$$L(t, x^h, \dot{x}^h) - L(t, x^h, \psi^h) + p_j \psi^j - \frac{\partial S}{\partial x^j} \dot{x}^j \geq 0,$$

or, if we use (2.39) and (2.22),

$$L(t, x^h, \dot{x}^h) - L(t, x^h, \psi^h) - \frac{\partial L(t, x^h, \psi^h)}{\partial \dot{x}^j}(\dot{x}^j - \psi^j) \geq 0. \tag{2.44}$$

It is remarkable that we have been able to eliminate all explicit reference to the function $S(t, x^h)$ from the formulation of the condition (2.35). In fact, this suggests that we define, for any given Lagrangian $L$ and any pair of directional arguments $(\dot{x}^h, \psi^h)$, the so-called *excess function of Weierstrass*:

$$E(t, x^h, \psi^h, \dot{x}^h) = L(t, x^h, \dot{x}^h) - L(t, x^h, \psi^h) - \frac{\partial L(t, x^h, \psi^h)}{\partial \dot{x}^j}(\dot{x}^j - \psi^j), \tag{2.45}$$

in terms of which (2.44) can be written in the form

$$E(t, x^h, \psi^h, \dot{x}^h) \geq 0. \tag{2.46}$$

It is clear that this condition depends entirely on the nature of the given Lagrangian $L(t, x^h, \dot{x}^h)$, which may or may not be such as to satisfy (2.46). *However, for all physical applications the essential features of the theory are contained in the Hamilton–Jacobi equation* (2.41) *augmented by the covariant vector field defined by means of* (2.39). If, in addition, condition (2.46) is satisfied, it is immediately evident from the construction of the geodesic field that the curves of the field (such as $\Gamma$) do in fact afford an extreme value (at least locally) to the fundamental integral, and accordingly (2.46) is often referred to as the *sufficiency condition of Weierstrass* for the geodesic field. For many physical systems the Lagrangian is such that this condition is not satisfied; nevertheless, the fundamental field equations (2.41) and (2.39) yield physically acceptable conclusions. It would seem, therefore, that the fundamental laws of nature are not, as is usually supposed, concerned with extreme values of a fundamental integral; instead, *these laws are derivable from the properties of fields defined entirely by means of the Hamilton–Jacobi equation*.

In conclusion we note that the second mean value theorem allows us to express the excess function (2.45) in a more compact form as follows:

$$E(t, x^h, \psi^h, \dot{x}^h) = \frac{1}{2} \frac{\partial^2 L(t, x^h, z^h)}{\partial \dot{x}^j\, \partial \dot{x}^k}(\dot{x}^j - \psi^j)(\dot{x}^k - \psi^k), \tag{2.47}$$

where

$$z^h = \theta \dot{x}^h + (1 - \theta)\psi^h \qquad 0 < \theta < 1. \tag{2.48}$$

For example, in the case of the Lagrangian (1.6), we have from (1.5) that

$$\frac{\partial^2 L}{\partial \dot{x}^j\, \partial \dot{x}^h} = a_{jh},$$

so that

$$E(t, x^h, \psi^h, \dot{x}^h) = \tfrac{1}{2} a_{jh}(\dot{x}^j - \psi^j)(\dot{x}^h - \psi^h),$$

and thus, since the $a_{jh}$ are the coefficients of a positive definite quadratic form, it follows that the Weierstrass condition (2.46) is automatically satisfied for dynamical systems whose Lagrangian is given by (1.6).

### EXAMPLE

One of the simplest and yet most illuminating examples of the above analysis is the theory of the classical harmonic oscillator. This model consists of a single particle (assumed to be of unit mass) moving along a straight line, the latter to be taken as our $x$-axis. The particle is attracted to the origin by a force $F = -\omega^2 x$, where $\omega$ is a given positive constant. The corresponding potential is therefore $V = \tfrac{1}{2}\omega^2 x^2$, while the kinetic energy is $T = \tfrac{1}{2}\dot{x}^2$, where $\dot{x} = dx/dt$, and $t$ denotes the time. According to (1.6) the Lagrangian of the system is given by

$$L = \tfrac{1}{2}(\dot{x})^2 - \tfrac{1}{2}\omega^2 x^2. \tag{2.49}$$

Clearly $n = 1$ in this example. The condition (2.24) is obviously satisfied, so that the canonical momentum may be defined as

$$p = \frac{\partial L}{\partial \dot{x}} = \dot{x}, \tag{2.50}$$

and thus the Hamiltonian function (2.25) assumes the form

$$H(x, p) = \tfrac{1}{2}p^2 + \tfrac{1}{2}\omega^2 x^2. \tag{2.51}$$

Since

$$\frac{\partial H}{\partial x} = \omega^2 x, \qquad \frac{\partial H}{\partial p} = p, \tag{2.52}$$

the canonical equations (2.43) and (2.26) are simply

$$\dot{p} = -\omega^2 x, \qquad \dot{x} = p. \tag{2.53}$$

Let us choose our initial conditions such that $x = A$ and $\dot{x} = 0$ when $t = 0$, where $A$ is some constant. Under these conditions the solution of (2.53) is given by

$$x = A \cos \omega t. \tag{2.54}$$

Thus $x$ lies in the range $[-A, A]$, and the motion is periodic with period $2\pi/\omega$ (or frequency $\omega/2\pi$) and has an amplitude $|A|$. In the configuration space $R_2$ of the variables $(x, t)$ the solution (2.54) is represented by a cosine

curve which intersects the $t$-axis at the points $t = (2k+1)\pi/2\omega$, $k = 0, \pm 1, \pm 2, \ldots$. A one-parameter family $F$ of such extremal curves is obtained when one allows $A$ to vary continuously from $-\infty$ to $\infty$. With the exception of the points $(0, (2k+1)\pi/2\omega)$ this family covers $R_2$ simply in the sense that through any point of $R_2$ [other than $(0, (2k+1)\pi/2\omega)$] there passes a unique member of $F$. The latter thus defines a geodesic field on $R_2$.

Along any member $\Gamma$ of $F$ we have, by (2.51) and (2.53),

$$\frac{dH}{dt} = p\dot{p} + \omega^2 x\dot{x} = p(-\omega^2 x) + \omega^2 xp = 0,$$

so that $H$ is constant along $\Gamma$, this value of $H$ being denoted by $E$. Clearly the value of $E$ depends upon the choice of $\Gamma$; that is, $E$ is a function of the parameter $A$. In fact, from (2.54) we have

$$p = \dot{x} = -A\omega \sin \omega t \tag{2.55}$$

and if this, together with (2.54), is substituted in (2.51), we obtain

$$E = \tfrac{1}{2}A^2\omega^2. \tag{2.56}$$

Let $P(t, x)$ be an arbitrary point (not on the $t$- or $x$-axes of $R_2$). The value $A$ of the parameter corresponding to the member $\Gamma$ of $F$ which passes through $P$ is given by (2.54), or

$$A = x \sec \omega t. \tag{2.57}$$

Because of (2.56) the value of $E$ corresponding to this point is

$$E = \tfrac{1}{2}\omega^2 x^2 \sec^2 \omega t, \tag{2.58}$$

while (2.55) and (2.57) yield

$$p = -\omega x \sec \omega t \sin \omega t = -\omega x \tan \omega t. \tag{2.59}$$

According to (2.41) and (2.51) the Hamilton–Jacobi equation associated with this example is

$$\frac{\partial S}{\partial t} + H\left(x, \frac{\partial S}{\partial x}\right) \equiv \frac{\partial S}{\partial t} + \frac{1}{2}\left(\frac{\partial S}{\partial x}\right)^2 + \frac{1}{2}\omega^2 x^2 = 0. \tag{2.60}$$

In order to construct a solution $S = S(t, x)$ of this equation we recall the relation (2.39) of our general theory, namely,

$$p = \frac{\partial S}{\partial x}, \tag{2.61}$$

in which we substitute from (2.59), obtaining

$$\frac{\partial S}{\partial x} = -\omega x \tan \omega t. \tag{2.62}$$

Moreover, since $H = E$ by construction, it follows from (2.58) and the left-hand side of (2.60) that

$$\frac{\partial S}{\partial t} = -\frac{1}{2} \omega^2 x^2 \sec^2 \omega t. \tag{2.63}$$

By differentiation of (2.62) and (2.63) with respect to $t$ and $x$, respectively, it is easily verified that the integrability conditions of this system are satisfied (as is to be expected from the general theory). Hence $S$ is found by direct integration of (2.62) and (2.63), which yields

$$S(t, x) = -\tfrac{1}{2}\omega x^2 \tan \omega t \tag{2.64}$$

(the constant of integration being neglected). This is the required solution of the Hamilton–Jacobi equation (2.60). The relations

$$S(t, x) = \Sigma, \tag{2.65}$$

where $\Sigma$ is a variable parameter, define a one-parameter family of curves in $R_2$, which is said to be *transversal* to the family $F$.

Along an extremal $\Gamma$ of the family $F$ corresponding to the parameter value $A$, the function (2.64) can, by means of (2.54), be expressed in the form

$$S = -\tfrac{1}{2}A^2\omega \sin \omega t \cos \omega t = -\tfrac{1}{4}A^2\omega \sin 2\omega t. \tag{2.66}$$

Similarly, along $\Gamma$ the Lagrangian (2.49) has the value

$$L = \tfrac{1}{2}A^2\omega^2(\sin^2 \omega t - \cos^2 \omega t) = -\tfrac{1}{2}A^2\omega^2 \cos 2\omega t. \tag{2.67}$$

Thus, along $\Gamma$,

$$\frac{dS}{dt} = L, \tag{2.68}$$

which is consistent with the relation (2.34) [subject to (2.31)] of the general theory. Moreover, between any two points on $\Gamma$ corresponding to the values $t = t_1, t = t_2$, chosen such that the interval $[t_1, t_2]$ does not contain a value $t = (2k+1)\pi/2\omega$ $(k = 0, \pm 1, \pm 2, \ldots)$ or zeros of (2.67), we have

$$\int_{\Gamma\, t_1}^{t_2} L \, dt = S(t_2) - S(t_1) = \tfrac{1}{4}A^2\omega[\sin 2\omega t_1 - \sin 2\omega t_2], \tag{2.69}$$

which clearly indicates that the function $S$ represents a measure of the value of the fundamental integral when evaluated along an extremal $\Gamma$ of the family $F$.

*Remark.*   In most textbooks on analytical mechanics (e.g., Goldstein [1], p. 278) the Hamilton–Jacobi equation (2.60) is treated in a somewhat different fashion. Noting that the Hamiltonian (2.51) does not depend

explicitly on the time, while $H = E$, a solution of the form

$$S = -Et + S^*(x), \tag{2.70}$$

is postulated. When this is substituted in (2.60) it is found with the aid of (2.56) that

$$S^* = \int \sqrt{2E - \omega^2 x^2} \, dx = \omega \int \sqrt{A^2 - x^2} \, \dot{x} \, dt, \tag{2.71}$$

so that

$$S = -\tfrac{1}{2}A^2\omega^2 t + \omega \int \sqrt{A^2 - x^2} \, \dot{x} \, dt. \tag{2.72}$$

This is obviously somewhat different from the solution (2.64) above, which is due to the following reason. Because of the introduction of $E$ in (2.70) one immediately restricts oneself to a particular member $\Gamma$ of the family $F$, namely, that member whose parameter value $A$ corresponds to $E$ in accordance with (2.56). Thus the solution (2.72) can only reflect the value of $S$ on $\Gamma$, instead of on the whole of $R_2$ (with the exception of isolated points). However, it is easily seen that (2.72) coincides with the special case (2.66) of (2.64). For it follows from the above argument that the integral on the right-hand side of (2.72) should be taken along $\Gamma$, which requires that we should substitute from (2.54), and this yields

$$S = \tfrac{1}{2}A^2\omega^2 \int_{\Gamma} (\sin^2 \omega t - \tfrac{1}{2}) \, dt = -\tfrac{1}{4}A^2\omega \sin 2\omega t, \tag{2.73}$$

where it should be noted that the negative square root of $A^2 - x^2$ has to be chosen in the course of this calculation in order to achieve consistency of (2.72), (2.73), and (2.61) with (2.55).

This remark should serve to emphasize the following important observation. The function $S$, which plays such a fundamental role in our general theory as a solution of the Hamilton–Jacobi equation, is a function of the variables $(t, x)$, and should be expressed as such, and should not be represented as a function of parameters such as $A$ which merely serve to label individual members of the field $F$. This observation applies to variational problems in general.

## 6.3   INVARIANCE PROPERTIES OF THE FUNDAMENTAL INTEGRAL: THE THEOREM OF NOETHER FOR SINGLE INTEGRALS

In Section 6.1 the invariance properties of the fundamental integral of a problem in the calculus of variations were discussed with reference to general coordinate transformations on our differentiable manifold $X_n$ on the one hand, and general parameter transformations on the other. It may

happen, however, that a given fundamental integral is invariant under a special group of continuous transformations—irrespective of coordinate or parameter transformations—and, under these circumstances, the corresponding Lagrangian must satisfy certain conditions. The latter can be expressed in a very concise manner in terms of the Euler–Lagrange vector $E_j(L)$, and the resulting formulation of these conditions is usually referred to as Noether's theorem (Noether [1]). The importance of this theorem is due to the fact that it allows us to construct quantities which are constant along any extremal, that is, a curve which satisfies the Euler–Lagrange equations $E_j(L) = 0$, and thus, in the case of physical applications, one obtains relations which may be interpreted as conservation laws.

Let us suppose that we are given some $r$-parameter group of transformations which is represented in the form

$$\bar{x}^j = \bar{x}^j(t, x^h, w^s) \qquad \bar{t} = \bar{t}(t, x^h, w^s), \tag{3.1}$$

in which the $w^s$ $(s = 1, \ldots, r)$ denote the $r$ parameters of the group (see, e.g., Eisenhart [2]). [In the sequel the superscripts $s, t$ range from 1 to $r$; the summation convention is operative also in respect of these indices.] It will be assumed that the functions on the right-hand sides of (3.1) are of class $C^3$ in all arguments, and that the values $w^s = 0$ $(s = 1, \ldots, r)$ correspond to the identity transformation $\bar{x}^j = x^j$, $\bar{t} = t$. Furthermore, the parameters $w^s$ are supposed to be entirely independent of each other. As a result of these stipulations it is possible to express (3.1) in the form

$$\bar{x}^j = x^j + \zeta_s^j(t, x^h)w^s + \frac{1}{2!} \zeta_s{}^j{}_t(t, x^h)w^s w^t + \cdots,$$
$$\bar{t} = t + \xi_s(t, x^h)w^s + \frac{1}{2!} \xi_{st}(t, x^h)w^s w^t + \cdots, \tag{3.2}$$

so that

$$\left(\frac{\partial \bar{x}^j}{\partial w^s}\right)_{w^s = 0} = \zeta_s^j, \qquad \left(\frac{\partial \bar{t}}{\partial w^s}\right)_{w^s = 0} = \xi_s. \tag{3.3}$$

It is also immediately evident from (3.2) that

$$\left(\frac{\partial \bar{x}^j}{\partial x^h}\right)_{w^s = 0} = \delta_h^j, \quad \left(\frac{\partial \bar{x}^j}{\partial t}\right)_{w^s = 0} = 0, \quad \left(\frac{\partial \bar{t}}{\partial x^h}\right)_{w^s = 0} = 0, \quad \left(\frac{\partial \bar{t}}{\partial t}\right)_{w^s = 0} = 1, \tag{3.4}$$

together with

$$\left(\frac{\partial^2 \bar{x}^j}{\partial w^s \partial x^h}\right)_{w^s = 0} = \frac{\partial \zeta_s^j}{\partial x^h}, \qquad \left(\frac{\partial^2 \bar{x}^j}{\partial w^s \partial t}\right)_{w^s = 0} = \frac{\partial \zeta_s^j}{\partial t}, \tag{3.5}$$

and

$$\left(\frac{\partial^2 \bar{t}}{\partial w^s \partial x^h}\right)_{w^s = 0} = \frac{\partial \xi_s}{\partial x^h}, \qquad \left(\frac{\partial^2 \bar{t}}{\partial w^s \partial t}\right)_{w^s = 0} = \frac{\partial \xi_s}{\partial t}. \tag{3.6}$$

Let us now assume that we are given a problem in the calculus of variations which is such that its fundamental integral is invariant under the transformation (3.1), that is, that

$$\int_C L\left(\bar{t}, \bar{x}^j, \frac{d\bar{x}^j}{d\bar{t}}\right) d\bar{t} = \int_C L\left(t, x^h, \frac{dx^h}{dt}\right) dt, \tag{3.7}$$

the integration being performed along an arbitrary curve $C$ of $X_n$. In order that this requirement be satisfied, it is necessary and sufficient that

$$L\left(\bar{t}, \bar{x}^j, \frac{d\bar{x}^j}{d\bar{t}}\right) \frac{d\bar{t}}{dt} = L\left(t, x^h, \frac{dx^h}{dt}\right) \tag{3.8}$$

for any set of values of the parameters $w^s$. Accordingly we may differentiate (3.8) with respect to $w^s$, noting that the right-hand side is independent of these parameters. Putting $w^s = 0$ after differentiation, we thus obtain

$$\left\{\frac{\partial L}{\partial t}\left(\frac{\partial \bar{t}}{\partial w^s}\right)_{w^s=0} + \frac{\partial L}{\partial x^j}\left(\frac{\partial \bar{x}^j}{\partial w^s}\right)_{w^s=0} + \frac{\partial L}{\partial \dot{x}^j}\left[\frac{\partial}{\partial w^s}\left(\frac{d\bar{x}^j}{d\bar{t}}\right)\right]_{w^s=0}\right\}\left(\frac{d\bar{t}}{dt}\right)_{w^s=0}$$
$$+ L\frac{\partial}{\partial w^s}\left(\frac{d\bar{t}}{dt}\right)_{w^s=0} = 0. \tag{3.9}$$

In order to be able to exploit this condition we must evaluate the various derivatives with respect to $w^s$ which appear on the left-hand side.

To this end we observe that, along any curve $C$, we have by (3.1)

$$\frac{d\bar{x}^j}{dt} = \frac{\partial \bar{x}^j}{\partial t} + \frac{\partial \bar{x}^j}{\partial x^h} \dot{x}^h, \tag{3.10}$$

and

$$\frac{d\bar{t}}{dt} = \frac{\partial \bar{t}}{\partial t} + \frac{\partial \bar{t}}{\partial x^h} \dot{x}^h, \tag{3.11}$$

in which $\dot{x}^h = dx^h/dt$ refers to the tangent vector of $C$. Relative to the $\bar{x}^j$-coordinates, and the parameter $\bar{t}$, this tangent vector is represented by $d\bar{x}^j/d\bar{t}$, which is given by

$$\frac{d\bar{x}^j}{d\bar{t}} \frac{d\bar{t}}{dt} = \frac{d\bar{x}^j}{dt}. \tag{3.12}$$

Thus

$$\frac{\partial}{\partial w^s}\left(\frac{d\bar{x}^j}{d\bar{t}}\right)\frac{d\bar{t}}{dt} + \frac{d\bar{x}^j}{d\bar{t}}\frac{\partial}{\partial w^s}\left(\frac{d\bar{t}}{dt}\right) = \frac{\partial}{\partial w^s}\left(\frac{d\bar{x}^j}{dt}\right),$$

or, if we substitute from (3.10) and (3.11),

$$\frac{\partial}{\partial w^s}\left(\frac{d\bar{x}^j}{d\bar{t}}\right)\frac{d\bar{t}}{dt} + \frac{d\bar{x}^j}{d\bar{t}}\left(\frac{\partial^2\bar{t}}{\partial w^s\,\partial t} + \frac{\partial^2\bar{t}}{\partial w^s\,\partial x^h}\dot{x}^h\right) = \frac{\partial^2\bar{x}^j}{\partial w^s\,\partial t} + \frac{\partial^2\bar{x}^j}{\partial w^s\,\partial x^h}\dot{x}^h. \tag{3.13}$$

In (3.10) and (3.11) we put $w^s = 0$, noting (3.4) at the same time. This gives

$$\left(\frac{d\bar{t}}{dt}\right)_{w^s=0} = 1, \qquad \left(\frac{d\bar{x}^j}{dt}\right)_{w^s=0} = \delta_h^j\dot{x}^h = \dot{x}^j, \tag{3.14}$$

so that (3.12) yields

$$\left(\frac{d\bar{x}^j}{d\bar{t}}\right)_{w^s=0} = \dot{x}^j. \tag{3.15}$$

We now put $w^s = 0$ in (3.13), and substitute from (3.5), (3.6), (3.14), and (3.15), obtaining

$$\begin{aligned}\frac{\partial}{\partial w^s}\left(\frac{d\bar{x}^j}{d\bar{t}}\right)_{w^s=0} &= \left(\frac{\partial\zeta_s^j}{\partial t} + \frac{\partial\zeta_s^j}{\partial x^h}\dot{x}^h\right) - \dot{x}^j\left(\frac{\partial\xi_s}{\partial t} + \frac{\partial\xi_s}{\partial x^h}\dot{x}^h\right)\\ &= \frac{d\zeta_s^j}{dt} - \dot{x}^j\frac{d\xi_s}{dt}.\end{aligned} \tag{3.16}$$

Also, it is immediately evident from (3.11) and (3.6) that

$$\frac{\partial}{\partial w^s}\left(\frac{d\bar{t}}{dt}\right)_{w^s=0} = \frac{\partial\xi_s}{\partial t} + \frac{\partial\xi_s}{\partial x^h}\dot{x}^h = \frac{d\xi_s}{dt}. \tag{3.17}$$

We now substitute from (3.3), (3.16), (3.14), and (3.17) in (3.9). This yields

$$\frac{\partial L}{\partial t}\xi_s + \frac{\partial L}{\partial x^j}\zeta_s^j + \frac{\partial L}{\partial\dot{x}^j}\left(\frac{d\zeta_s^j}{dt} - \dot{x}^j\frac{d\xi_s}{dt}\right) + L\frac{d\xi_s}{dt} = 0. \tag{3.18}$$

This, then, is a necessary condition which must be satisfied as a result of our invariance requirement (3.7). We shall refer to (3.18) as an *invariance identity*: it is clearly a condition involving the Lagrangian $L$ and the quantities $\zeta_s^j$, $\xi_s$ which are specified uniquely by the transformation group (3.1) under which the fundamental integral is to be invariant. It should be observed that the condition (3.18) does not involve the second or higher derivatives of the functions $x^j(t)$ which define the arbitrary curve $C$.

The theorem of Noether can be derived directly from (3.18) when the Euler-Lagrange vector $E_j(L)$ is introduced into this relation. To this end we note that

$$\frac{\partial L}{\partial t}\xi_s = \left(\frac{dL}{dt} - \frac{\partial L}{\partial x^j}\dot{x}^j - \frac{\partial L}{\partial\dot{x}^j}\ddot{x}^j\right)\xi_s,$$

and

$$\frac{\partial L}{\partial\dot{x}^j}\frac{d\zeta_s^j}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^j}\zeta_s^j\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^j}\right)\zeta_s^j,$$

together with

$$\frac{\partial L}{\partial\dot{x}^j}\dot{x}^j\frac{d\xi_s}{dt} + \frac{\partial L}{\partial\dot{x}^j}\ddot{x}^j\xi_s = \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^j}\dot{x}^j\xi_s\right) - \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^j}\right)\dot{x}^j\xi_s.$$

These three relations are substituted in turn in (3.18). After a little simplification this yields

$$\begin{aligned}&\left[\frac{\partial L}{\partial x^j} - \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^j}\right)\right]\zeta_s^j - \left[\frac{\partial L}{\partial x^j} - \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}^j}\right)\right]\dot{x}^j\xi_s\\ &\qquad\qquad + \frac{d}{dt}\left[L\xi_s + \frac{\partial L}{\partial\dot{x}^j}\zeta_s^j - \frac{\partial L}{\partial\dot{x}^j}\dot{x}^j\xi_s\right] = 0,\end{aligned}$$

or, if we invoke the definition (1.18) of the Euler-Lagrange vector,

$$E_j(L)(\zeta_s^j - \dot{x}^j\xi_s) = \frac{d}{dt}\left[L\xi_s + \frac{\partial L}{\partial\dot{x}^j}(\zeta_s^j - \dot{x}^j\xi_s)\right]. \tag{3.19}$$

This relation represents $r$ identities ($s = 1, \ldots, r$); its implications can be enunciated in the following form:

*If the fundamental integral of a problem in the calculus of variations is invariant under a given r-parameter group of transformations such as (3.1), then there exist r distinct linear combinations of the Euler-Lagrange vector $E_j(L)$ which are exact differentials.*

The right-hand side of (3.19) suggests that we introduce the notation

$$\theta_s = -L\xi_s - \frac{\partial L}{\partial\dot{x}^j}(\zeta_s^j - \dot{x}^j\xi_s). \tag{3.20}$$

By means of (2.22) and (2.25) these quantities can be expressed as follows in terms of the canonical coordinates:

$$\theta_s = H\xi_s - p_j\zeta_s^j, \tag{3.21}$$

and (3.19) can be written in the form

$$E_j(L)(\zeta_s^j - \dot{x}^j\xi_s) = -\frac{d\theta_s}{dt}. \tag{3.22}$$

This conclusion immediately gives rise to the usual formulation of

## NOETHER'S THEOREM

*If the fundamental integral of a problem in the calculus of variations is invariant under the r-parameter group of transformations* (3.1), *the r distinct quantities $\theta_s$ defined by* (3.20) *or* (3.21) *are constant along any extremal.*

The following examples should serve to illustrate typical applications of this theorem.

### EXAMPLE 1

Let us suppose that the fundamental integral is invariant under the one-parameter group (e.g., a time translation)

$$\bar{x}^j = x^j, \qquad \bar{t} = t + w.$$

According to (3.3) we have $\zeta^j = 0$, $\xi = 1$ in this case. From (3.18) it is immediately evident that

$$\frac{\partial L}{\partial t} = 0, \tag{3.23}$$

while (3.21) gives

$$\theta = H.$$

The theorem then asserts that $H = \text{const.}$ along any extremal. This, incidentally, can also be verified directly as follows:

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial x^j} \dot{x}^j + \frac{\partial H}{\partial p_j} \frac{dp_j}{dt},$$

or, if we use (2.26), (2.28), and (2.29),

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t} + \dot{x}^j \left( \frac{\partial H}{\partial x^j} + \frac{dp_j}{dt} \right) = -\frac{\partial L}{\partial t} + E_j(L)\dot{x}^j,$$

so that $dH/dt = 0$ along any extremal by virtue of (3.23).

### EXAMPLE 2

Let us suppose that the fundamental integral is invariant under the *n*-parameter transformation group (e.g., a space translation)

$$\bar{x}^j = x^j + w^j, \qquad \bar{t} = t \qquad (j = 1, \ldots, n = r).$$

According to (3.3) we have $\zeta_h^j = \delta_h^j$, $\xi_j = 0$ in this case. From (3.21) we then have

$$\theta_j = -p_j,$$

and the theorem asserts that $p_j = \text{const.}$ along any extremal. Again this is easily verified by means of (3.18), from which it is inferred that $\partial L/\partial x^j = 0$ under these circumstances.

If, in the above examples, the Lagrangian $L$ refers to a dynamical system of $n$ degrees of freedom, one would interpret the above conclusions as the conservation laws of energy and linear momentum, respectively.

## 6.4  INTEGRAL INVARIANTS AND THE INDEPENDENT HILBERT INTEGRAL

Fields of extremal curves possess many striking geometrical properties, some of which are of considerable importance in many applications of the calculus of variations. This is true in particular of geometrical optics, in which, according to Fermat's principle, bundles of light rays are represented by such fields (Carathéodory [1], Synge [2]). The geometrical properties to be discussed here are based on the concept of the so-called integral invariants (E. Cartan [2], Hilbert [1]), which had originally been introduced by Poincaré primarily in connection with the theory of first-order ordinary differential equations.

Let us consider an *n*-parameter family of curves in $X_{n+1}$, represented parametrically by $n$ class $C^2$ functions in the form

$$x^j = x^j(t, u^h) \qquad (j, h, \ldots, = 1, \ldots, n), \tag{4.1}$$

in which the $u^h$ denote the parameters of the family. It is assumed that

$$\frac{\partial(x^1, \ldots, x^n)}{\partial(u^1, \ldots, u^n)} \neq 0, \tag{4.2}$$

so that the family (4.1) covers a region $R$ of $X_{n+1}$ simply in the sense that through each point $P \in R$ there passes one and only one curve of (4.1); henceforth our considerations will be restricted to this region. In accordance with our previous notation we shall write

$$\dot{x}^j = \frac{\partial x^j}{\partial t} \tag{4.3}$$

along each member of (4.1), and, by means of the definition (2.22), we associate the covariant canonical momentum vector field $p_j$ with the contravariant vector field (4.3), namely,

$$p_j = p_j(t, u^h). \tag{4.4}$$

Corresponding to an arbitrary displacement $(dt, dx^j)$ at any point $P(t, x^h)$ of $R$ we can now construct the 1-form

$$\omega = p_j \, dx^j - H(t, x^h, p_h) \, dt, \tag{4.5}$$

which, as we shall now see, possesses some very remarkable properties. We shall begin by evaluating the exterior derivative of this 1-form. According

to the theory of Section 5.2 we have (recalling that $dt \wedge dt = 0$)

$$d\omega = dp_j \wedge dx^j - dH \wedge dt = \left( dp_j + \frac{\partial H}{\partial x^j} dt \right) \wedge dx^j - \frac{\partial H}{\partial p_j} dp_j \wedge dt. \qquad (4.6)$$

But from (4.1) and (4.4) we have, in terms of the notation (4.3),

$$dx^j = \dot{x}^j \, dt + \frac{\partial x^j}{\partial u^h} du^h, \quad dp_j = \dot{p}_j \, dt + \frac{\partial p_j}{\partial u^k} du^k, \qquad (4.7)$$

where $\dot{p}_j = \partial p_j / \partial t$, so that (4.6) is equivalent to

$$d\omega = \left[ \left( \dot{p}_j + \frac{\partial H}{\partial x^j} \right) dt + \frac{\partial p_j}{\partial u^k} du^k \right] \wedge \left( \dot{x}^j \, dt + \frac{\partial x^j}{\partial u^h} du^h \right) - \frac{\partial H}{\partial p_j} \frac{\partial p_j}{\partial u^k} du^k \wedge dt$$

$$= \left( \dot{p}_j + \frac{\partial H}{\partial x^j} \right) \frac{\partial x^j}{\partial u^h} dt \wedge du^h + \left( \dot{x}^j - \frac{\partial H}{\partial p_j} \right) \frac{\partial p_j}{\partial u^h} du^h \wedge dt$$

$$- \frac{\partial x^j}{\partial u^h} \frac{\partial p_j}{\partial u^k} du^h \wedge du^k. \qquad (4.8)$$

Because of the identity (2.26) the coefficient of $\partial p_j / \partial u^h$ vanishes identically. Moreover, the last expression on the right-hand side suggests the introduction of the so-called *Lagrange brackets*, which are defined as

$$[u^h, u^k] = \frac{\partial x^j}{\partial u^h} \frac{\partial p_j}{\partial u^k} - \frac{\partial x^j}{\partial u^k} \frac{\partial p_j}{\partial u^h}. \qquad (4.9)$$

It should be remarked that these quantities play a central role in analytical mechanics and geometrical optics (Goldstein [1], Whittaker [1]). By virtue of the skew-symmetry of $du^h \wedge du^k$ in the indices $h$ and $k$ we have

$$\frac{\partial x^j}{\partial u^h} \frac{\partial p_j}{\partial u^k} du^h \wedge du^k = \tfrac{1}{2} [u^h, u^k] \, du^h \wedge du^k,$$

and thus (4.8) can be expressed in the form

$$d\omega = \left( \dot{p}_j + \frac{\partial H}{\partial x^j} \right) \frac{\partial x^j}{\partial u^h} dt \wedge du^h - \tfrac{1}{2} [u^h, u^k] \, du^h \wedge du^k, \qquad (4.10)$$

which is the expression we have been seeking.

Now let us suppose that the curves of the family (4.1) are *extremals* of our variational problem: that is, it is assumed that these curves satisfy the canonical equations (2.43):

$$\dot{p}_j + \frac{\partial H}{\partial x^j} = 0. \qquad (4.11)$$

[However, it is not assumed at this stage that these extremals are necessarily members of a geodesic field in the sense that there exists a function $S(t, x^h)$ such that $p_j$ and $H$ are given as derivatives of $S$ as in (2.39) and (2.41), respectively.] Under these conditions the expression (4.10) reduces to

$$d\omega = -\tfrac{1}{2} [u^h, u^k] \, du^h \wedge du^k. \qquad (4.12)$$

When we substitute from (4.1) and (4.4) in the definition (4.9) of our Lagrange brackets, we see that the latter may be expressed as functions of the $(n + 1)$ variables $(u^h, t)$. Thus, applying the lemma of Poincaré [namely, $d(d\omega) = 0$] to (4.12), we obtain

$$0 = \frac{\partial}{\partial t} [u^h, u^k] \, dt \wedge du^h \wedge du^k + \frac{\partial}{\partial u^l} [u^h, u^k] \, du^l \wedge du^h \wedge du^k. \qquad (4.13)$$

But it follows directly from (4.9) that

$$\frac{\partial}{\partial u^l} [u^h, u^k] = \frac{\partial^2 x^j}{\partial u^l \partial u^h} \frac{\partial p_j}{\partial u^k} + \frac{\partial x^j}{\partial u^h} \frac{\partial^2 p_j}{\partial u^l \partial u^k} - \frac{\partial^2 x^j}{\partial u^l \partial u^k} \frac{\partial p_j}{\partial u^h} - \frac{\partial x^j}{\partial u^k} \frac{\partial^2 p_j}{\partial u^l \partial u^h},$$

in which each term is symmetric in some pair of the triplet $l, h, k$. Hence

$$\frac{\partial}{\partial u^l} [u^h, u^k] \, du^l \wedge du^h \wedge du^k = 0 \qquad (4.14)$$

identically, and accordingly we infer from (4.13) that

$$\frac{\partial}{\partial t} [u^h, u^k] = 0. \qquad (4.15)$$

This shows that *the Lagrange brackets are constant along each member of any family of extremals.* This theorem was first established by Lagrange on the basis of some very complicated analysis; it subsequently played a significant role in the development of the theory of optical instruments.

From among the $n$ parameters $u^h$ we now single out two variables, say, $u^1$ and $u^2$, and we construct the two-dimensional plane $T_2$ defined by the equations $u^3 = a^3, \ldots, u^n = a^n$, in the domain $R_n$ of the $u^h$, where $a^3, \ldots, a^n$ are arbitrary constants. Let $g$ denote a region on $T_2$ which is bounded by a simple closed curve $\partial g$, of which it is assumed that it may be represented parametrically in the form $u^1 = u^1(\tau)$, $u^2 = u^2(\tau)$, these functions being periodic and of class $C^2$ in $\tau$. According to (4.12) the 2-form $d\omega$ assumes the value

$$d\omega = -[u^1, u^2] \, du^1 \wedge du^2 \qquad (4.16)$$

on $g$. An application of Stoke's theorem then yields

$$- \int_g [u^1, u^2] \, du^1 \wedge du^2 = \int_g d\omega = \int_{\partial g} \omega.$$

But, by virtue of (4.15), the integrand on the left-hand side is independent of $t$, and therefore *the integral on the right is simply a constant whose value depends solely on the choice of the curve $\partial g$*:

$$\int_{\partial g} \omega = \kappa(\partial g). \tag{4.17}$$

It is a simple matter to give a geometrical interpretation of this result directly in terms of the family (4.1) of extremals in $X_{n+1}$. To each value of the parameter $\tau$, that is, to each point on $\partial g$, there corresponds a unique extremal of the family (4.1), which is given by

$$x^j = x^j(t, u^1(\tau), u^2(\tau), a^3, \ldots, a^n), \tag{4.18}$$

so that the totality of points on $\partial g$ defines a one-parameter subfamily of (4.1) (which may be pictured as the generators of the two-dimensional surface of a tube in $X_{n+1}$). Let $c$ be any closed curve in $X_{n+1}$ which intersects each member of (4.18) once and only once (that is, any simple curve which encircles the aforementioned tube). Clearly every curve of this kind is an image of the curve $\partial g$ on $T_2$ under the map defined by (4.18), and it therefore follows from (4.5) and (4.17) that

$$\oint_c (p_j \, dx^j - H \, dt) = \kappa(\partial g). \tag{4.19}$$

*Thus the value of this integral depends solely on the choice of the one-parameter family (4.18), and is the same for all closed curves which intersect each member of this family precisely once.*

There are two important special cases of this conclusion which we now consider. First, let $c_1$ and $c_2$ represent two closed curves of the type specified above, but subject to the restriction that they be constrained to lie on the hypersurfaces $t = t_1$ and $t = t_2$ of $X_{n+1}$, respectively. From (4.19) it is inferred that for each of these curves we have

$$\int_{c_1} p_j \, dx^j = \kappa(\partial g), \qquad \int_{c_2} p_j \, dx^j = \kappa(\partial g), \tag{4.20}$$

despite the fact that the integrands refer to entirely different values of $t$. It therefore follows that *the value of the integral*

$$\int_{c_1} p_j \, dx^j \tag{4.21}$$

*is independent of $t$.* An integral possessing this property is called an integral invariant (or, more precisely, a *relative* integral invariant, since the domain of integration is closed).

Second, let us now suppose that the 1-form (4.5) is exact, that is, that there exists a function $S(t, x^h)$ such that

$$\omega = dS. \tag{4.22}$$

Under these circumstances we clearly have

$$p_j = \frac{\partial S}{\partial x^j}, \qquad \frac{\partial S}{\partial t} + H = 0, \tag{4.23}$$

and from the theory of Section 6.2 it follows that our family (4.1) of extremals constitutes a geodesic field. Moreover, we now infer from (4.22) that

$$d\omega = 0, \tag{4.24}$$

and hence, by virtue of the converse of Poincaré's lemma, we conclude that *the condition (4.24) is characteristic of geodesic fields.* Equivalently, such fields are characterized among arbitrary families of extremals by the condition that all Lagrange brackets vanish, as is evident immediately from (4.12). The constant $\kappa(\partial g)$ obviously vanishes identically for any one-parameter subfamily of an $n$-parameter family of extremals of a geodesic field. More generally, for *any* differentiable curve $\bar{C}: \bar{x}^j = \bar{x}^j(t)$ of $X_{n+1}$, with $\dot{\bar{x}}^j = d\bar{x}^j/dt$, which joins two points $P_1(t_1, \bar{x}^j(t_1))$, $P_2(t_2, \bar{x}^j(t_2))$, we have from (4.23) and (4.22)

$$J \equiv \int_{\bar{C}\,t_1}^{t_2} \{p_j \dot{\bar{x}}^j - H(t, x^h, p_h)\} \, dt = \int_{\bar{C}\,P_1}^{P_2} \omega = S(P_2) - S(P_1), \tag{4.25}$$

from which it is evident that the value of the integral $J$ depends solely on the end-points $P_1$ and $P_2$ and not on the choice of the curve $\bar{C}$ joining these points. Indeed, the relation (4.25) defines the so-called *independent integral of Hilbert* (Hilbert [1]). It should be emphasized, however, that the arguments $p_h$ in the integrand of the latter refer to the geodesic field. By means of (2.22) and (2.25) the canonical variables may be eliminated, so that (4.25) can be represented in the following equivalent form:

$$J = \int_{\bar{C}\,t_1}^{t_2} \left\{ L(t, x^h, \dot{x}^h) + \frac{\partial L(t, x^h, \dot{x}^h)}{\partial \dot{x}^j} (\dot{\bar{x}}^j - \dot{x}^j) \right\} dt, \tag{4.26}$$

where it should be remarked once more that the arguments $\dot{x}^j$ in the above integrand refer to the geodesic field. Needless to say, the Hilbert independent integral is defined if and only if a geodesic field exists.

As a special case of (4.26) let us suppose that the points $P_1$, $P_2$ on $\bar{C}$ are chosen such that a given extremal $\Gamma$ of the geodesic field passes through $P_1$ and $P_2$. If, for the moment, we identify $\bar{C}$ with $\Gamma$, we may put $\dot{\bar{x}}^j = \dot{x}^j$ in

(4.26), and the latter reduces to

$$J = \int_{\Gamma}^{t_2}_{t_1} L(t, x^h, \dot{x}^h)\, dt. \tag{4.27}$$

It therefore follows from (4.26) and (4.27) that, for an *arbitrary* differentiable curve $\bar{C}$ joining $P_1$ and $P_2$, we have

$$\int_{\bar{C}}^{t_2}_{t_1} L(t, x^h, \dot{x}^h)\, dt - \int_{\Gamma}^{t_2}_{t_1} L(t, x^h, \dot{x}^h)\, dt$$

$$= \int_{\bar{C}}^{t_2}_{t_1} \left\{ L(t, x^h, \dot{\bar{x}}^h) - L(t, x^h, \dot{x}^h) - \frac{\partial L(t, x^h, \dot{x}^h)}{\partial \dot{x}^j}(\dot{\bar{x}}^j - \dot{x}^j) \right\} dt, \tag{4.28}$$

or, if we invoke the definition (2.45) of the Weierstrass excess function,

$$\int_{\bar{C}}^{t_2}_{t_1} L(t, x^h, \dot{\bar{x}}^h)\, dt - \int_{\Gamma}^{t_2}_{t_1} L(t, x^h, \dot{x}^h)\, dt = \int_{\bar{C}}^{t_2}_{t_1} E(t, x^h, \dot{x}^h, \dot{\bar{x}}^h)\, dt. \tag{4.29}$$

This result is sometimes referred to as the *fundamental formula of the calculus of variations*: it indicates quite clearly that the extremal $\Gamma$ affords a minimum to the fundamental integral relative to all curves $\bar{C}$ with the same end-points $P_1$, $P_2$ whenever the Weierstrass condition $E(t, x^h, \dot{x}^h, \dot{\bar{x}}^h) \geq 0$ is satisfied. [Strictly speaking, certain additional requirements must be imposed in order to ensure that the above conclusion is always valid: namely, that the points $P_1$, $P_2$ are sufficiently close together to exclude the possibility of the existence of a so-called conjugate point of $P_1$ on $\Gamma$ between $P_1$ and $P_2$, and that $\bar{C}$ be contained in a certain neighborhood of $\Gamma$; however, a precise description of these conditions is beyond the scope of the present discussion.]

Before concluding this section, let us return to the construction of integral invariants of higher order by means of a family of extremals (it being no longer assumed that the latter is necessarily associated with a geodesic field). To this end we shall now restrict our attention to a hypersurface $t = t_1 = \text{const.}$ of $X_{n+1}$, on which the 1-form (4.5) is given by $\omega = p_j\, dx^j$. For some integer $p$, $1 \leq p \leq n$, let us consider the $2p$-form defined by

$$(d\omega)^p = d\omega \wedge \cdots \wedge d\omega \qquad (p \text{ exterior products}). \tag{4.30}$$

Now, let $R_{2n}$ denote the $2n$-dimensional *phase space*, whose points are represented by the $2n$ coordinates $(x^j, p_j)$. Over some simply-connected $2p$-dimensional region $h_{2p}$ of $R_{2n}$, which is bounded by a closed, class $C^1$ subspace $\partial h_{2p}$ of dimension $2p - 1$, we construct the integral

$$K_{2p} = \int_{h_{2p}} (d\omega)^p. \tag{4.31}$$

However, according to (4.12) and (4.15) the 2-form $d\omega$ is independent of the value $t = t_1$, and this is true therefore also of the integral (4.31). Thus the latter is called an *integral invariant of order* $2p$ (or, more precisely, an *absolute* integral invariant, since the region $h_{2p}$ is essentially arbitrary). Moreover, since

$$(d\omega)^p = d[\omega \wedge (d\omega)^{p-1}], \tag{4.32}$$

it follows from Stokes's theorem that

$$K_{2p} = \int_{h_{2p}} d[\omega \wedge (d\omega)^{p-1}] = \int_{\partial h_{2p}} \omega \wedge (d\omega)^{p-1}, \tag{4.33}$$

from which we conclude that the integral on the right-hand side is a *relative integral invariant of order* $2p - 1$. [For $p = 1$, the latter obviously corresponds to (4.21).]

Since $d\omega = dp_j \wedge dx^j$ on the hypersurface $t = t_1$ of $X_{n+1}$, it follows from (4.30) that

$$(d\omega)^p = (-1)^{\frac{1}{2}p(p-1)} dp_{j_1} \wedge \cdots \wedge dp_{j_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p}, \tag{4.34}$$

because a total of $1 + 2 + \cdots + (p - 1) = \frac{1}{2}p(p - 1)$ interchanges of the 1-forms $dp_j$ is required to write the product (4.30) in the order given in (4.34). Thus the integral invariant (4.31) of order $2p$ may also be expressed in the form

$$K_{2p} = (-1)^{\frac{1}{2}p(p-1)} \int_{h_{2p}} dp_{j_1} \wedge \cdots \wedge dp_{j_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_p}. \tag{4.35}$$

In particular, for $p = n$, we have from (4.34)

$$(d\omega)^n = (-1)^{\frac{1}{2}n(n-1)} \varepsilon_{j_1 \cdots j_n} \varepsilon^{j_1 \cdots j_n} dp_1 \wedge \cdots \wedge dp_n \wedge dx^1 \wedge \cdots \wedge dx^n,$$

or, if we use (4.2.22)

$$(d\omega)^n = (-1)^{\frac{1}{2}n(n-1)} n!\, dp_1 \wedge \cdots \wedge dp_n \wedge dx^1 \wedge \cdots \wedge dx^n. \tag{4.36}$$

It therefore follows from (4.31), with $p = n$, that the $2n$-fold integral

$$\int_{h_{2n}} dp_1 \wedge \cdots \wedge dp_n \wedge dx^1 \wedge \cdots \wedge dx^n \tag{4.37}$$

is an absolute integral invariant of order $2n$. This is *the famous theorem of Liouville* which is of fundamental importance in statistical mechanics, (see, e.g., Uhlenbeck and Ford [1]).

## 6.5  MULTIPLE INTEGRAL PROBLEMS IN THE CALCULUS OF VARIATIONS

The theory of the preceding sections can be generalized to the case when the variational problem is concerned with extreme values of a multiple integral. This extension will be briefly discussed here (Carathéodory [2]).

Instead of a single parameter $t$ we now consider a set of $m$ parameters $t^\alpha$. (Here, and in the sequel, all Greek indices range from 1 to $m$; the summation convention is operative also in respect of these indices.) The $n$ dependent variables are again denoted by $x^j$. It is advisable to introduce a new configuration space $X_{n+m}$, which is supposed to be an $(n + m)$-dimensional manifold whose local coordinates may be identified with the $n + m$ variables $t^\alpha, x^j$.

An $m$-dimensional subspace $C_m$ of $X_{n+m}$ can be represented parametrically in the form:

$$x^j = x^j(t^\alpha). \tag{5.1}$$

It will be supposed that the functions $x^j(t^\alpha)$ are of class $C^2$; one may therefore construct the derivatives

$$\dot{x}_\alpha^j = \frac{\partial x^j}{\partial t^\alpha}, \tag{5.2}$$

and our attention will be restricted to subspaces for which the rank of the matrix $(\dot{x}_\alpha^j)$ is $m$. Under a class $C^2$ coordinate transformation of the type

$$\bar{x}^j = \bar{x}^j(x^h), \tag{5.3}$$

with nonvanishing Jacobian, involving solely the variables $x^h$, the derivatives (5.2) transform as follows:

$$\dot{\bar{x}}_\alpha^j = \frac{\partial \bar{x}^j}{\partial t^\alpha} = \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial x^h}{\partial t^\alpha} = \frac{\partial \bar{x}^j}{\partial x^h} \dot{x}_\alpha^h, \tag{5.4}$$

from which it is evident that the $\dot{x}_\alpha^h$ represent the components of a contravariant vector relative to the transformation (5.3). However, under a class $C^2$ parameter transformation of the type

$$\bar{t}^\alpha = \bar{t}^\alpha(t^\beta), \tag{5.5}$$

with nonvanishing Jacobian, which involves solely the parameters $t^\beta$, one has

$$\frac{\partial x^j}{\partial \bar{t}^\beta} = \frac{\partial x^j}{\partial t^\alpha} \frac{\partial t^\alpha}{\partial \bar{t}^\beta} = \frac{\partial t^\alpha}{\partial \bar{t}^\beta} \dot{x}_\alpha^j, \tag{5.6}$$

from which it follows that the derivatives $\dot{x}_\alpha^j$ are the components of a covariant vector relative to the transformation (5.5).

On a given subspace $C_m$ any function $f = f(t^\alpha, x^j, \dot{x}_\alpha^j)$ can be expressed entirely as a function of $t^\alpha$ by substitution of $x^j$ and $\dot{x}_\alpha^j$ from (5.1) and (5.2). The following notation will be used consistently:

$$\frac{df}{dt^\alpha} = \frac{\partial f}{\partial t^\alpha} + \frac{\partial f}{\partial x^j} \dot{x}_\alpha^j + \frac{\partial f}{\partial \dot{x}_\beta^j} \ddot{x}_{\beta}{}^j{}_\alpha, \tag{5.7}$$

where we have written

$$\ddot{x}_{\beta}{}^j{}_\alpha = \frac{\partial^2 x^j}{\partial t^\beta \, \partial t^\alpha} = \ddot{x}_{\alpha}{}^j{}_\beta. \tag{5.8}$$

Now let us suppose that we are given a function $L(t^\alpha, x^j, \dot{x}_\alpha^j)$, which is supposed to be of class $C^2$ in its $(m + n + mn)$ arguments. Then, for a prescribed closed, simply-connected region $G$ in the domain of the variables $t^\alpha$, one can construct the $m$-fold integral

$$I(C_m) = \int_G L(t^\alpha, x^j, \dot{x}_\alpha^j) \, d(t), \tag{5.9}$$

where we have used the notation

$$d(t) = dt^1 \, dt^2 \cdots dt^m. \tag{5.10}$$

The value of this integral generally depends on the subspace $C_m$ by means of which it is defined, for it is to be understood that the arguments $x^j, \dot{x}_\alpha^j$ in the integrand $L$ are to be replaced by the functions (5.1) and (5.2) of the independent variables $t^\alpha$ which are specified by $C_m$. Again we are concerned with extreme values of the integral (5.9). More precisely, we seek the conditions which a subspace $C_m$ must satisfy in order that it afford an extreme value to this integral relative to the values which it will assume for all other class $C^2$ subspaces $\tilde{C}_m$ which coincide with $C_m$ on the boundary $\partial G$ of $G$ (see Figure 7). Thus any admissible subspace of comparison $\tilde{C}_m$, which is represented in the form $x^j = \tilde{x}^j(t^\alpha)$, is supposed to satisfy the restriction

$$\tilde{x}^j(t^\alpha) = x^j(t^\alpha) = a^j(t^\alpha) \quad \text{for all } t^\alpha \in \partial G, \tag{5.11}$$

where the functions $a^j(t^\alpha)$ are fixed, prescribed functions on the boundary $\partial G$ of $G$. (This condition corresponds to the stipulation of the single integral problem according to which the admissible curves of comparison are required to terminate at two fixed end-points $P_1$ and $P_2$; as in the single integral case this restriction may be relaxed in various ways.)

As before, we refer to (5.9) as the fundamental integral of our problem in the calculus of variations, whose integrand is again called the Lagrangian.

Before turning to the variational problem as such, let us briefly glance at the invariance conditions which may be imposed on the integral (5.9).
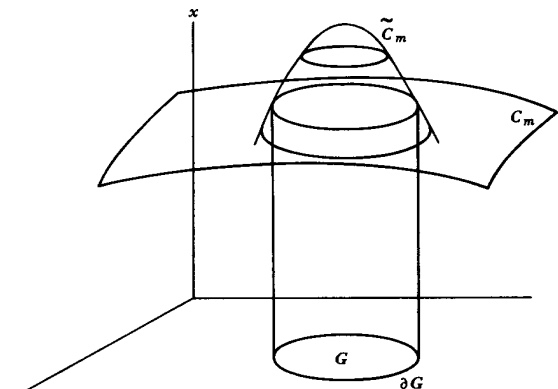
Domain of variables $t^\alpha$

Fig. 7

First, we shall demand that (5.9) be invariant under the coordinate transformations (5.3). This obviously implies that *the Lagrangian L be a scalar relative to* (5.3). Second, under the inverse of the parameter transformation (5.5), the integral (5.9) assumes the form

$$I(C_m) = \int_G L\left\{t^\alpha(\bar{t}^\beta), x^j[t^\alpha(\bar{t}^\beta)], \frac{\partial x^j}{\partial \bar{t}^\beta} \frac{\partial \bar{t}^\beta}{\partial t^\alpha}\right\} B \, d(\bar{t}), \qquad (5.12)$$

in which

$$B = \det\left(\frac{\partial t^\alpha}{\partial \bar{t}^\beta}\right) \qquad (5.13)$$

denotes the Jacobian of the inverse of (5.5). The integral (5.9) is said to be parameter-invariant if

$$I(C_m) = \int_G L\left(\bar{t}^\alpha, x^j, \frac{\partial x^j}{\partial \bar{t}^\alpha}\right) d(\bar{t}) \qquad (5.14)$$

for all parameter transformations of the type (5.5). The necessary and sufficient conditions which the Lagrangian L must satisfy in order that (5.9) be parameter-invariant are contained in the following relations:

$$\frac{\partial L}{\partial t^\alpha} = 0, \qquad (5.15)$$

together with

$$\frac{\partial L}{\partial \dot{x}_\alpha^j} \dot{x}_\beta^j = L\delta_\beta^\alpha. \qquad (5.16)$$

In the immediate sequel it is not assumed that (5.9) is parameter-invariant, so that no appeal to (5.15) and (5.16) will be made. Thus we shall not give a proof of these conditions here; instead, we shall deduce the latter from the theorem of Noether for multiple integrals which is to be established presently.

However, since L is supposed to be a scalar under the coordinate transformation (5.3), we have

$$\bar{L}(t^\alpha, \bar{x}^j, \dot{\bar{x}}_\alpha^j) = L(t^\beta, x^k, \dot{x}_\beta^k), \qquad (5.17)$$

so that

$$\frac{\partial L}{\partial \dot{x}_\beta^k} = \frac{\partial \bar{L}}{\partial \dot{\bar{x}}_\alpha^j} \frac{\partial \dot{\bar{x}}_\alpha^j}{\partial \dot{x}_\beta^k}, \qquad (5.18)$$

and

$$\frac{\partial L}{\partial x^k} = \frac{\partial \bar{L}}{\partial \bar{x}^j} \frac{\partial \bar{x}^j}{\partial x^k} + \frac{\partial \bar{L}}{\partial \dot{\bar{x}}_\alpha^j} \frac{\partial \dot{\bar{x}}_\alpha^j}{\partial x^k}. \qquad (5.19)$$

But from (5.4) it follows directly that

$$\frac{\partial \dot{\bar{x}}_\alpha^j}{\partial \dot{x}_\beta^k} = \frac{\partial \bar{x}^j}{\partial x^h} \frac{\partial \dot{x}_\alpha^h}{\partial \dot{x}_\beta^k} = \frac{\partial \bar{x}^j}{\partial x^h} \delta_k^h \delta_\alpha^\beta = \frac{\partial \bar{x}^j}{\partial x^k} \delta_\alpha^\beta, \qquad (5.20)$$

together with

$$\frac{\partial \dot{\bar{x}}_\alpha^j}{\partial x^k} = \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \dot{x}_\alpha^h. \qquad (5.21)$$

When (5.20) is substituted in (5.18) one obtains

$$\frac{\partial L}{\partial \dot{x}_\beta^k} = \frac{\partial \dot{x}^j}{\partial x^k} \delta_\alpha^\beta \frac{\partial \bar{L}}{\partial \dot{\bar{x}}_\alpha^j} = \frac{\partial \bar{x}^j}{\partial x^k} \frac{\partial \bar{L}}{\partial \dot{\bar{x}}_\beta^j}, \qquad (5.22)$$

from which it follows that these quantities represent the components of a covariant vector. However, when (5.21) is substituted in (5.19) it is found that

$$\frac{\partial L}{\partial x^k} = \frac{\partial \bar{L}}{\partial \bar{x}^j} \frac{\partial \bar{x}^j}{\partial x^k} + \frac{\partial \bar{L}}{\partial \dot{\bar{x}}_\alpha^j} \frac{\partial^2 \bar{x}^j}{\partial x^h \partial x^k} \dot{x}_\alpha^h, \qquad (5.23)$$

which indicates that the derivatives $\partial L/\partial x^k$ are nontensorial. In order to augment these quantities in a suitable manner, we differentiate (5.22) with respect to $t^\beta$ (summing over $\beta$). In terms of the notation defined in (5.7)

we thus find that

$$\frac{d}{dt^\beta}\left(\frac{\partial L}{\partial \dot{x}^k_\beta}\right) = \frac{\partial \bar{x}^j}{\partial x^k}\frac{d}{dt^\beta}\left(\frac{\partial \bar{L}}{\partial \dot{\bar{x}}^j_\beta}\right) + \frac{\partial \bar{L}}{\partial \dot{\bar{x}}^j_\beta}\frac{\partial^2 \bar{x}^j}{\partial x^h\,\partial x^k}\dot{x}^h_\beta.$$

The second term on the right-hand side of this relation is seen to be identical with the corresponding term in (5.23) after the repeated indices $\beta$ have been replaced by $\alpha$. On eliminating these terms it is found that

$$\frac{d}{dt^\beta}\left(\frac{\partial L}{\partial \dot{x}^k_\beta}\right) - \frac{\partial L}{\partial x^k} = \frac{\partial \bar{x}^j}{\partial x^k}\left[\frac{d}{dt^\beta}\left(\frac{\partial \bar{L}}{\partial \dot{\bar{x}}^j_\beta}\right) - \frac{\partial \bar{L}}{\partial \bar{x}^j}\right],$$

from which we infer that the quantities defined by

$$E_j(L) = \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right) - \frac{\partial L}{\partial x^j}, \qquad (5.24)$$

constitute the components of a covariant vector, which will again be referred to as the *Euler–Lagrange vector*.

We shall now give a brief discussion of the multiple integral variational problem as formulated above. Again several distinct methods are feasible. According to the classical procedure one can evaluate an expression for the first variation $\delta I$ of the integral $I$; if the subspace $C_m$ is to afford an extreme value to $I$ it is necessary that $\delta I = 0$, and it may be shown—by a method analogous to that sketched in Section 6.2—that this implies that the functions $x^j = x^j(t^\alpha)$ which define the subspace $C_m$ must then satisfy the *system of n second-order partial differential equations*

$$E_j(L) = 0. \qquad (5.25)$$

These are the so-called *Euler–Lagrange equations*: they represent necessary conditions which the subspace $C_m$ has to satisfy in order that it afford an extreme value to the fundamental integral (5.9).

An important example of a multiple integral variational problem with $n = 1$ and $m$ arbitrary is represented by the Lagrangian

$$L = \tfrac{1}{2}a^{\alpha\beta}\dot{x}_\alpha \dot{x}_\beta + \tfrac{1}{2}\mu^2 x^2,$$

in which $a^{\alpha\beta} = a^{\beta\alpha}$ and $\mu$ are assumed to be constants, while $\dot{x}_\alpha = \partial x/\partial t^\alpha$. Since $\partial L/\partial \dot{x}_\alpha = a^{\alpha\beta}\dot{x}_\beta$, $\partial L/\partial x = \mu^2 x$, it follows from (5.24) that the Euler–Lagrange equations (5.25) assume the form

$$a^{\alpha\beta}\frac{\partial^2 x}{\partial t^\alpha\,\partial t^\beta} = \mu^2 x.$$

When $m = 4$ and $a^{\alpha\beta} = \delta^{\alpha\beta}$, this is the famous Klein–Gordon equation of a scalar field for a suitably chosen value of $\mu$; moreover, when $m = 3$ and

$a^{\alpha\beta} = \delta^{\alpha\beta}$, and $\mu = 0$, the above equation is simply Laplace's equation (see, e.g., Davis [1]).

Instead of outlining the procedure based on the first variation here, we briefly discuss the corresponding generalization of the method of equivalent integrals. This entails the construction of a function $\Delta(t^\alpha, x^j, \dot{x}^j_\alpha)$ which is the integrand of an independent integral in the sense that this integral does not depend on the choice of the subspace $C_m$ from the class of subspaces satisfying condition (5.11); an integrand of this kind obviously represents a counterpart of the total derivative $dS/dt$ as introduced in Section 6.2 for the case of a single integral. Such a function $\Delta$ can be obtained as follows.

Let $S^\alpha = S^\alpha(t^\varepsilon, x^h)$ denote a set of $m$ class $C^2$ functions. Relative to a subspace $C_m$ as defined by (5.1) one can construct the quantities

$$c^\alpha_\beta = \frac{dS^\alpha}{dt^\beta} = \frac{\partial S^\alpha}{\partial t^\beta} + \frac{\partial S^\alpha}{\partial x^j}\dot{x}^j_\beta, \qquad (5.26)$$

which define a determinant

$$\Delta(t^\varepsilon, x^h, \dot{x}^h_\varepsilon) = \det(c^\alpha_\beta), \qquad (5.27)$$

in which the cofactors of the elements $c^\alpha_\beta$ are denoted by $C^\beta_\alpha$; that is,

$$C^\alpha_\beta c^\beta_\varepsilon = \delta^\alpha_\varepsilon \Delta. \qquad (5.28)$$

A single equation of the type $V(t^\varepsilon, x^h) = 0$ represents an $(n + m - 1)$-dimensional subspace of $X_{n+m}$; thus the system of $m$ equations

$$S^\alpha(t^\varepsilon, x^h) = \Sigma^\alpha, \qquad (5.29)$$

in which the $\Sigma^\alpha$ denote $m$ variable parameters, represent an $m$-parameter family of subspaces of dimension $m + n - m = n$. It will be assumed that this family covers a region of $X_{n+m}$ simply, that is, through each point of this region there passes one and only one member of the family, and our subsequent considerations are restricted to this region. The intersection of the family (5.29) with some subspace $C_m: x^j = x^j(t^\alpha)$ is given by

$$S^\alpha(t^\varepsilon, x^h(t^\varepsilon)) = \Sigma^\alpha. \qquad (5.30)$$

Thus for a given value $t^\alpha_0 \in G$ there corresponds a point $(t^\alpha_0, x^h(t^\alpha_0))$ on $C_m$, through which there passes a member of (5.29) corresponding to parameter values $\Sigma^\alpha_0$, the latter being given by $\Sigma^\alpha_0 = S^\alpha(t^\varepsilon_0, x^h(t^\varepsilon_0))$. In this sense, therefore, the equation (5.30) defines a mapping of the region $G$ of the domain $X_m$ of the variables $t^\alpha$ onto a region $G_\Sigma$ of the domain $R_m$ of the parameters $\Sigma^\alpha$. It will be assumed henceforth that $C_m$ is such that the Jacobian of the mapping be positive, that is, that

$$\frac{\partial(\Sigma^1, \dots, \Sigma^m)}{\partial(t^1, \dots, t^m)} = \det\left(\frac{d\Sigma^\alpha}{dt^\beta}\right) = \det\left(\frac{dS^\alpha}{dt^\beta}\right) = \det(c^\alpha_\beta) = \Delta > 0, \qquad (5.31)$$

which implies that the mapping (5.30) is locally one-to-one. Moreover, from the general transformation formulae for definite integrals we infer that

$$\int_G \Delta \, d(t) = \int_G \frac{\partial(\Sigma^1, \ldots, \Sigma^m)}{\partial(t^1, \ldots, t^m)} \, d(t) = \int_{G_\Sigma} d\Sigma^1 \cdots d\Sigma^m. \tag{5.32}$$

Now let us consider a second subspace $\tilde{C}_m : x^j = \tilde{x}^j(t^\alpha)$ satisfying the condition (5.11). Analogously to (5.30) this also defines a mapping of $X_m$ into $R_m$, namely, by means of the equations

$$S^\alpha(t^\varepsilon, \tilde{x}^h(t^\varepsilon)) = \Sigma^\alpha, \tag{5.33}$$

it being assumed once more that the corresponding Jacobian $\tilde{\Delta}$ is positive. Furthermore, because of (5.11),

$$S^\alpha(t^\varepsilon, x^h(t^\varepsilon)) = S^\alpha(t^\varepsilon, \tilde{x}^h(t^\varepsilon)) \quad \text{for all } t^\alpha \in \partial G,$$

so that the images of the boundary $\partial G$ of $G$ under the two transformations (5.30) and (5.33) coincide. It follows that the image of $G$ under (5.33) is again $G$, and hence, as in (5.32),

$$\int_G \tilde{\Delta} \, d(t) = \int_{G_\Sigma} d\Sigma^1 \cdots d\Sigma^m,$$

so that

$$\int_G \tilde{\Delta} \, d(t) = \int_G \Delta \, d(t) \tag{5.34}$$

for all subspaces $\tilde{C}_m$ satisfying the boundary condition (5.11). Thus the determinant $\Delta(t^\alpha, x^h, \dot{x}^h_\alpha)$ represents the integrand of the independent integral which we have been seeking.

For a given positive Lagrangian $L(t^\alpha, x^h, \dot{x}^h_\alpha)$ we can now construct the function

$$L^*(t^\alpha, x^h, \dot{x}^h_\alpha) = L(t^\alpha, x^h, \dot{x}^h_\alpha) - \Delta(t^\alpha, x^h, \dot{x}^h_\alpha), \tag{5.35}$$

which, by virtue of (5.34), is the Lagrangian of an *equivalent problem* in the sense that if $C_m$ affords an extreme value to $\int_G L^* \, d(t)$ it will do so also for $\int_G L \, d(t)$ (and conversely) relative to all admissible subspaces of comparison satisfying (5.11). In order to exploit this situation let us suppose that it is possible to construct $m$ functions $S^\alpha(t^\varepsilon, x^h)$ together with $nm$ functions $\psi^j_\alpha(t^\varepsilon, x^h)$ which are such that

$$L^*(t^\alpha, x^h, \dot{x}^h_\alpha) = 0 \quad \text{whenever } \dot{x}^j_\alpha = \psi^j_\alpha, \tag{5.36}$$

while

$$L^*(t^\alpha, x^h, \dot{x}^h_\alpha) \geq 0 \quad \text{whenever } \dot{x}^j_\alpha \neq \psi^j_\alpha. \tag{5.37}$$

If the subspace $C_m : x^j = x^j(t^\alpha)$ is a solution of $\dot{x}^j_\alpha = \psi^j_\alpha(t^\varepsilon, x^h)$, subject to (5.11), the conditions (5.36), (5.37) obviously imply that $C_m$ affords a minimum to $\int_G L^* \, d(t)$ relative to all subspaces satisfying (5.11), and hence, in view of the remarks made above, the subspace $C_m$ represents a solution to our variational problem.

The functions $S^\alpha(t^\varepsilon, x^h)$, $\psi^j_\alpha(t^\varepsilon, x^h)$ on which this construction is based are said to constitute a *geodesic field*, whose properties we shall now investigate. It is evident from (5.36) and (5.37) that $L^*(t^\alpha, x^h, \dot{x}^h_\alpha)$ assumes a minimum value when $\dot{x}^j_\alpha = \psi^j_\alpha$ (for $t^\alpha$, $x^h$ fixed); it is therefore necessary that $\partial L^*/\partial \dot{x}^j_\alpha = 0$ for these values. Because of (5.35) this is equivalent to the statement that

$$\frac{\partial L}{\partial \dot{x}^j_\alpha} = \frac{\partial \Delta}{\partial \dot{x}^j_\alpha} \tag{5.38}$$

for $\dot{x}^j_\alpha = \psi^j_\alpha$, or, because of the general formula (1.3.15) for the derivative of a determinant,

$$\frac{\partial L}{\partial \dot{x}^j_\alpha} = C^\beta_\varepsilon \frac{\partial c^\varepsilon_\beta}{\partial \dot{x}^j_\alpha}.$$

But from (5.26) it follows that

$$\frac{\partial c^\varepsilon_\beta}{\partial \dot{x}^j_\alpha} = \frac{\partial S^\varepsilon}{\partial x^h} \frac{\partial \dot{x}^h_\beta}{\partial \dot{x}^j_\alpha} = \frac{\partial S^\varepsilon}{\partial x^h} \delta^h_j \delta^\alpha_\beta = \frac{\partial S^\varepsilon}{\partial x^j} \delta^\alpha_\beta, \tag{5.39}$$

so that

$$\frac{\partial L}{\partial \dot{x}^j_\alpha} = C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial x^j} \tag{5.40}$$

for $\dot{x}^j_\alpha = \psi^j_\alpha$. Moreover, by means of (5.35), (5.28), and (5.26) the condition (5.36) can be expressed as

$$L\delta^\alpha_\beta = \Delta\delta^\alpha_\beta = C^\alpha_\varepsilon c^\varepsilon_\beta = C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial x^j} \dot{x}^j_\beta, \tag{5.41}$$

or, if we invoke (5.40),

$$L\delta^\alpha_\beta = C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta. \tag{5.42}$$

This suggests the introduction of the so-called *Hamiltonian complex*

$$H^\alpha_\beta = -L\delta^\alpha_\beta + \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta, \tag{5.43}$$

which allows us to express (5.42) in the form

$$H^\alpha_\beta = -C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta}. \tag{5.44}$$

The equations (5.40) and (5.44) characterize the geodesic field; however, from the construction (5.41) it is evident that the $m^2$ relations (5.44) are not independent, being equivalent to the single statement (5.36), which is

$$L(t^\alpha, x^h, \dot{x}^h_\alpha) = \Delta(t^\alpha, x^h, \dot{x}^h_\alpha) \tag{5.45}$$

for $\dot{x}^j_\alpha = \psi^j_\alpha$. It should also be noted that, for the case $m = 1$, the relations (5.40) and (5.44) respectively reduce to the equations (2.39) and (2.41) when the canonical variables in the latter are replaced by means of their corresponding counterparts as functions of $(t, x^j, \dot{x}^j)$.

Now, let us differentiate (5.28) with respect to $t^\alpha$:

$$\frac{dC^\alpha_\beta}{dt^\alpha} c^\beta_\varepsilon + C^\alpha_\beta \frac{dc^\beta_\varepsilon}{dt^\alpha} = \delta^\alpha_\varepsilon \frac{d\Delta}{dt^\alpha} = \frac{d\Delta}{dt^\varepsilon}. \tag{5.46}$$

But from the rule (1.3.15) and (5.26) we have

$$\frac{d\Delta}{dt^\varepsilon} = C^\alpha_\beta \frac{dc^\beta_\alpha}{dt^\varepsilon} = C^\alpha_\beta \frac{d^2 S^\beta}{dt^\varepsilon\, dt^\alpha} = C^\alpha_\beta \frac{d^2 S^\beta}{dt^\alpha\, dt^\varepsilon} = C^\alpha_\beta \frac{dc^\beta_\varepsilon}{dt^\alpha},$$

so that (5.46) gives

$$\frac{dC^\alpha_\beta}{dt^\alpha} c^\beta_\varepsilon = 0.$$

Since $\Delta = \det(c^\beta_\varepsilon) > 0$ by (5.31), this obviously gives rise to the identity

$$\frac{dC^\alpha_\beta}{dt^\alpha} = 0. \tag{5.47}$$

Thus differentiation of (5.40) with respect to $t^\alpha$ yields, in view of (5.26) and (5.47),

$$\frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right) = C^\alpha_\varepsilon \frac{d}{dt^\alpha}\left(\frac{\partial S^\varepsilon}{\partial x^j}\right) = C^\alpha_\varepsilon \left(\frac{\partial^2 S^\varepsilon}{\partial t^\alpha\, \partial x^j} + \frac{\partial^2 S^\varepsilon}{\partial x^h\, \partial x^j} \dot{x}^h_\alpha\right)$$

$$= C^\alpha_\varepsilon \frac{\partial}{\partial x^j}\left(\frac{\partial S^\varepsilon}{\partial t^\alpha} + \frac{\partial S^\varepsilon}{\partial x^h} \dot{x}^h_\alpha\right) = C^\alpha_\varepsilon \frac{\partial c^\varepsilon_\alpha}{\partial x^j} = \frac{\partial \Delta}{\partial x^j}. \tag{5.48}$$

On the other hand, when (5.45) is differentiated with respect to $x^j$, one obtains

$$\frac{\partial L}{\partial x^j} + \frac{\partial L}{\partial \dot{x}^h_\alpha} \frac{\partial \psi^h_\alpha}{\partial x^j} = \frac{\partial \Delta}{\partial x^j} + \frac{\partial \Delta}{\partial \dot{x}^h_\alpha} \frac{\partial \psi^h_\alpha}{\partial x^j},$$

or, because of (5.38),

$$\frac{\partial L}{\partial x^j} = \frac{\partial \Delta}{\partial x^j} \tag{5.49}$$

for the geodesic field. Hence (5.48) reduces to

$$\frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right) - \frac{\partial L}{\partial x^j} = 0, \tag{5.50}$$

from which we infer that *any subspace $C_m$, defined by means of the geodesic field as a solution $x^j = x^j(t^\alpha)$ of the equations $\dot{x}^j_\alpha = \psi^j_\alpha$, satisfies the Euler–Lagrange equations* (5.25). Such subspaces will again be called *extremals*.

Written out in full, the equation (5.50) assumes the following form:

$$\frac{\partial^2 L}{\partial \dot{x}^j_\alpha\, \partial \dot{x}^h_\beta} \frac{\partial^2 x^h}{\partial t^\alpha\, \partial t^\beta} + \frac{\partial^2 L}{\partial \dot{x}^j_\alpha\, \partial x^h} \frac{\partial x^h}{\partial t^\alpha} + \frac{\partial^2 L}{\partial \dot{x}^j_\alpha\, \partial t^\alpha} - \frac{\partial L}{\partial x^j} = 0. \tag{5.51}$$

As remarked above, this represents a system of $n$ second-order *partial* differential equations—in contrast to the single integral case, for which the Euler–Lagrange equations are ordinary differential equations. This phenomenon accounts for some of the most basic differences between single and multiple integral problems in the calculus of variations.

The following conclusion is of vital importance to all applications of the theory of multiple integral problems in the calculus of variations. When (5.43) is differentiated with respect to $t^\alpha$ one obtains

$$\frac{dH^\alpha_\beta}{dt^\alpha} = -\frac{dL}{dt^\beta} + \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right)\dot{x}^j_\beta + \frac{\partial L}{\partial \dot{x}^j_\alpha} \ddot{x}^j_{\alpha\,\beta}$$

$$= -\frac{\partial L}{\partial t^\beta} - \frac{\partial L}{\partial x^j}\dot{x}^j_\beta - \frac{\partial L}{\partial \dot{x}^j_\alpha}\ddot{x}^j_{\alpha\,\beta} + \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right)\dot{x}^j_\beta + \frac{\partial L}{\partial \dot{x}^j_\alpha}\ddot{x}^j_{\alpha\,\beta},$$

in which the third and the fifth terms on the right-hand side cancel, and thus, in terms of the notation (5.24),

$$\frac{dH^\alpha_\beta}{dt^\alpha} = -\frac{\partial L}{\partial t^\beta} + E_j(L)\dot{x}^j_\beta. \tag{5.52}$$

This is an identity; however, *for any extremal subspace $C_m$ it follows that*

$$\frac{dH^\alpha_\beta}{dt^\alpha} = -\frac{\partial L}{\partial t^\beta}. \tag{5.53}$$

Thus, whenever the Lagrangian $L(t^\alpha, x^h, \dot{x}^h_\alpha)$ does not depend explicitly on the independent variables $t^\alpha$, the *$m$ divergences $dH^\alpha_\beta/dt^\alpha$ vanish on an extremal subspace*. In many physical applications this gives rise directly to conservation laws.

At this stage it should be observed that the above theory of the geodesic field is based essentially on (5.36) and (5.38), the latter being a necessary consequence of the inequality (5.37). On the other hand, this inequality is

not necessarily implied by (5.36) and (5.38), and, because its validity is crucial to the method of equivalent integrals when one actually seeks to minimize the fundamental integral, it is desirable that we should investigate this state of affairs a little more closely. We henceforth let $\dot{x}_\alpha^j = \psi_\alpha^j$ refer to the geodesic field, while $\mathring{x}_\alpha^j$ is supposed to represent an arbitrary directional field. In terms of (5.35) the inequality (5.37) is simply

$$L(t^\alpha, x^h, \mathring{x}_\alpha^h) - \Delta(t^\alpha, x^h, \mathring{x}_\alpha^h) \geq 0, \tag{5.54}$$

and, in order to express this in a more useful form, we have to evaluate the determinant on the left-hand side. To this end we note that, by virtue of (5.28) and (5.31), the relations (5.40) and (5.44) are equivalent to

$$\frac{\partial S^\alpha}{\partial x^j} = \Delta^{-1} c_\varepsilon^\alpha \frac{\partial L}{\partial \dot{x}_\varepsilon^j}, \quad \frac{\partial S^\alpha}{\partial t^\beta} = -\Delta^{-1} c_\varepsilon^\alpha H_\beta^\varepsilon. \tag{5.55}$$

We may therefore write

$$\Delta(t^\alpha, x^h, \mathring{x}_\alpha^h) = \det\left(\frac{\partial S^\alpha}{\partial t^\beta} + \frac{\partial S^\alpha}{\partial x^j} \mathring{x}_\beta^j\right) = \det\left[\Delta^{-1} c_\varepsilon^\alpha\left(-H_\beta^\varepsilon + \frac{\partial L}{\partial \dot{x}_\varepsilon^j} \mathring{x}_\beta^j\right)\right]$$

$$= \Delta^{-m} \det(c_\varepsilon^\alpha)\det\left(-H_\beta^\varepsilon + \frac{\partial L}{\partial \dot{x}_\varepsilon^j} \mathring{x}_\beta^j\right),$$

where it should be clearly understood that $\Delta$, $H_\beta^\varepsilon$, $\partial L/\partial \dot{x}_\varepsilon^j$ refer to the geodesic field since these quantities were inserted by means of (5.55) which is valid solely for such fields. By means of (5.27), (5.43), and (5.45) this result can be expressed in the form

$$\Delta(t^\alpha, x^h, \mathring{x}_\alpha^h) = L^{1-m} \det\left[L\delta_\beta^\alpha + \frac{\partial L}{\partial \dot{x}_\alpha^j}(\mathring{x}_\beta^j - \dot{x}_\beta^j)\right]. \tag{5.56}$$

Thus the inequality (5.54) can be written as

$$E(t^\alpha, x^h, \dot{x}_\alpha^h, \mathring{x}_\alpha^h) \geq 0, \tag{5.57}$$

where we have defined the *Weierstrass excess function* by

$$E(t^\varepsilon, x^h, \dot{x}_\varepsilon^h, \mathring{x}_\varepsilon^h) = L(t^\varepsilon, x^h, \mathring{x}_\varepsilon^h)$$

$$- L^{1-m}(t^\varepsilon, x^h, \dot{x}_\varepsilon^h)\det\left[L(t^\varepsilon, x^h, \dot{x}_\varepsilon^h)\delta_\beta^\alpha + \frac{\partial L(t^\varepsilon, x^h, \dot{x}_\varepsilon^h)}{\partial \dot{x}_\alpha^j}(\mathring{x}_\beta^j - \dot{x}_\beta^j)\right]. \tag{5.58}$$

It is remarkable that we have succeeded in eliminating all explicit reference to the functions $S^\alpha$ in the final formulation (5.57) of the inequality (5.37). However, it should be stressed once more that, in the definition (5.58), the arguments $\dot{x}_\varepsilon^h$ refer to the geodesic field, while the $\mathring{x}_\varepsilon^j$ are arbitrary. The inequality (5.57) is usually called the *condition of Weierstrass*.

*Remark 1.* The above discussion of the multiple integral variational problem is a somewhat cursory one, many open questions having been ignored. A more comprehensive treatment would have to take into account various circumstances which are beyond the scope of this chapter. The most important of these is the fact that the function $\Delta$ as defined by (5.27) is not the only integrand which gives rise to an equivalent integral. For instance, it can be shown that each of the $m$ distinct functions $\Psi_{(r)}, r = 1, \ldots, m$, defined by

$$\Psi_{(r)}(t^\alpha, x^j, \dot{x}_\alpha^j) = \text{sum of all principal } r \times r \text{ minors of } \det\left(\frac{\partial S^\alpha}{\partial t^\beta} + \frac{\partial S^\alpha}{\partial x^j} \dot{x}_\beta^j\right), \tag{5.59}$$

represent integrands of distinct independent integrals, the function $\Delta$ corresponding to the special case $r = m$. It follows that each of the cases $r = 1, 2, \ldots, m$ gives rise to a different type of geodesic field: thus one has, in fact, a total of $m$ distinct field theories for $m$-fold integral problems in the calculus of variations. For each of these the Weierstrass excess function assumes a distinct form. Naturally all these distinct excess functions reduce to the function (2.45) when $m = 1$.

*Remark 2.* For each of these $m$ field theories a corresponding canonical formalism may be developed. It is for this reason that no canonical formalism is used in the above discussion; only a special canonical formalism peculiar to the independent integral (5.34) would have been appropriate. It is, in fact, possible to define a suitable Hamiltonian function, which in turn gives rise to a single Hamilton–Jacobi equation to be satisfied by the $m$ functions $S^\alpha(t^\varepsilon, x^h)$.

*Remark 3.* In the course of the construction of a subspace $C_m$ for which the vector field $\dot{x}_\alpha^j$ is identified with the vectors $\psi_\alpha^j(t^\varepsilon, x^h)$ of the geodesic field certain integrability conditions are implicitly involved. This is due to the fact that this identification is tantamount to the assumption that there exist solutions $x^j = x^j(t^\alpha)$ of the following system of partial differential equations:

$$\frac{\partial x^j}{\partial t^\alpha} = \psi_\alpha^j(t^\varepsilon, x^h), \tag{5.60}$$

which, according to the theory of Section 5.4, is possible if and only if the integrability conditions

$$\frac{\partial \psi_\alpha^j}{\partial t^\beta} + \frac{\partial \psi_\alpha^j}{\partial x^h} \psi_\beta^h = \frac{\partial \psi_\beta^j}{\partial t^\alpha} + \frac{\partial \psi_\beta^j}{\partial x^h} \psi_\alpha^h \tag{5.61}$$

are satisfied. For the classical problem in the calculus of variations described here these very restrictive additional conditions are unavoidable; it is, however, possible to construct satisfactory field theories for nonintegrable solution manifolds $C_m$.

## 6.6  THE THEOREM OF NOETHER FOR MULTIPLE INTEGRALS

It is possible to generalize the results derived in Section 6.3 for single integrals to the case of multiple integral problems in the calculus of variations (Noether [1], Rund [8]). Accordingly we shall consider an $r$-parameter group of transformations of the type

$$\bar{x}^j = \bar{x}^j(t^\beta, x^h, w^s), \qquad \bar{t}^\alpha = \bar{t}^\alpha(t^\beta, x^h, w^s), \tag{6.1}$$

in which the $w^s$ denote the $r$ independent parameters. (Here, and in the sequel, $s, t = 1, \ldots, r$; the summation convention applies to these indices also.) It is assumed that the functions on the right-hand sides of (6.1) are of class $C^3$, and that the values $w^1 = w^2 = \cdots = w^r = 0$ yield the identity transformation. Thus (6.1) can be expressed in the form

$$\left. \begin{aligned} \bar{x}^j &= x^j + \zeta_s^j(t^\beta, x^h)w^s + \frac{1}{2!}\zeta_{s\,t}^{\;j}(t^\beta, x^h)w^s w^t + \cdots, \\ \bar{t}^\alpha &= t^\alpha + \xi_s^\alpha(t^\beta, x^h)w^s + \frac{1}{2!}\xi_{s\,t}^{\;\alpha}(t^\beta, x^h)w^s w^t + \cdots, \end{aligned} \right\} \tag{6.2}$$

so that

$$\left(\frac{\partial \bar{x}^j}{\partial w^s}\right)_{w^s=0} = \zeta_s^j, \qquad \left(\frac{\partial \bar{t}^\alpha}{\partial w^s}\right)_{w^s=0} = \xi_s^\alpha. \tag{6.3}$$

Differentiation of (6.2) also yields the following relations, which will be used below:

$$\left(\frac{\partial \bar{x}^j}{\partial x^h}\right)_{w^s=0} = \delta_h^j, \qquad \left(\frac{\partial \bar{x}^j}{\partial t^\alpha}\right)_{w^s=0} = 0,$$

$$\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta}\right)_{w^s=0} = \delta_\beta^\alpha, \qquad \left(\frac{\partial \bar{t}^\alpha}{\partial x^h}\right)_{w^s=0} = 0; \tag{6.4}$$

$$\left(\frac{\partial^2 \bar{x}^j}{\partial w^s \, \partial t^\beta}\right)_{w^s=0} = \frac{\partial \zeta_s^j}{\partial t^\beta}, \qquad \left(\frac{\partial^2 \bar{x}^j}{\partial w^s \, \partial x^h}\right)_{w^s=0} = \frac{\partial \zeta_s^j}{\partial x^h}; \tag{6.5}$$

$$\left(\frac{\partial^2 \bar{t}^\alpha}{\partial w^s \, \partial t^\beta}\right)_{w^s=0} = \frac{\partial \xi_s^\alpha}{\partial t^\beta}, \qquad \left(\frac{\partial^2 \bar{t}^\alpha}{\partial w^s \, \partial x^h}\right)_{w^s=0} = \frac{\partial \xi_s^\alpha}{\partial x^h}. \tag{6.6}$$

Let us now consider some subspace $C_m$ of class $C^2$ represented parametrically by the equations

$$x^j = f^j(t^\alpha). \tag{6.7}$$

[For the purposes of the present section this notation is used instead of (5.1) in order to avoid possible confusion with the transformation equations (6.1).] In this context, then, the functions $\dot{x}_\alpha^j$ are given by

$$\dot{x}_\alpha^j = \frac{\partial f^j}{\partial t^\alpha}. \tag{6.8}$$

Relative to this subspace we have

$$\frac{d\bar{x}^j}{dt^\beta} = \frac{\partial \bar{x}^j}{\partial t^\beta} + \frac{\partial \bar{x}^j}{\partial x^h}\dot{x}_\beta^h, \tag{6.9}$$

together with

$$\frac{d\bar{t}^\alpha}{dt^\beta} = \frac{\partial \bar{t}^\alpha}{\partial t^\beta} + \frac{\partial \bar{t}^\alpha}{\partial x^h}\dot{x}_\beta^h. \tag{6.10}$$

From (6.4) it follows that

$$\left(\frac{d\bar{x}^j}{dt^\beta}\right)_{w^s=0} = \delta_h^j \dot{x}_\beta^h = \dot{x}_\beta^j, \tag{6.11}$$

while

$$\left(\frac{d\bar{t}^\alpha}{dt^\beta}\right)_{w^s=0} = \delta_\beta^\alpha. \tag{6.12}$$

In particular, for the functional determinant

$$D = \det\left(\frac{d\bar{t}^\alpha}{dt^\beta}\right), \tag{6.13}$$

it is evident that

$$(D)_{w^s=0} = 1. \tag{6.14}$$

In terms of the $(\bar{t}^\alpha, \bar{x}^j)$-coordinates the subspace $C_m$ as defined by (6.7) is supposed to be given in the form

$$\bar{x}^j = \bar{f}^j(\bar{t}^\alpha), \tag{6.15}$$

and, in analogy with (6.8), we shall write

$$\dot{\bar{x}}_\alpha^j = \frac{\partial \bar{f}^j}{\partial \bar{t}^\alpha}. \tag{6.16}$$

By definition, the functions (6.7) and (6.15) are related by means of the transformation group (6.1) according to the relation

$$\bar{x}^j(t^\beta, f^h(t^\beta), w^s) = \bar{f}^j[\bar{t}^\alpha(t^\beta, f^h(t^\beta), w^s)]. \tag{6.17}$$

Differentiation with respect to $t^\beta$ yields

$$\frac{\partial \bar{x}^j}{\partial t^\beta} + \frac{\partial \bar{x}^j}{\partial x^h}\frac{\partial f^h}{\partial t^\beta} = \frac{\partial \bar{f}^j}{\partial \bar{t}^\alpha}\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta} + \frac{\partial \bar{t}^\alpha}{\partial x^h}\frac{\partial f^h}{\partial t^\beta}\right),$$

or, if we use the notations (6.8) and (6.16),

$$\frac{\partial \bar{x}^j}{\partial t^\beta} + \frac{\partial \bar{x}^j}{\partial x^h}\dot{x}_\beta^h = \check{x}_\alpha^j\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta} + \frac{\partial \bar{t}^\alpha}{\partial x^h}\dot{x}_\beta^h\right). \tag{6.18}$$

By means of (6.4) it is inferred that, for $w^s = 0$,

$$\delta_h^j \dot{x}_\beta^h = (\check{x}_\alpha^j)_{w^s = 0}\delta_\beta^\alpha,$$

or

$$(\check{x}_\alpha^j)_{w^s = 0} = \dot{x}_\alpha^j. \tag{6.19}$$

Let us now assume that the fundamental integral (5.9) is invariant under the given $r$-parameter group (6.1). A necessary and sufficient condition for this to be the case is that

$$L(\bar{t}^\alpha, \bar{x}^j, \check{x}_\alpha^j)D = L(t^\beta, x^h, \dot{x}_\beta^h), \tag{6.20}$$

where it is to be clearly understood that $\check{x}_\alpha^j, \dot{x}_\beta^h$ are defined according to (6.8) and (6.16), respectively, while $D$ is given by (6.13).

The condition (6.20) holds for all values of the parameters $w^s$. We therefore differentiate partially with respect to $w^s$, after which we put $w^s = 0$, noting that the right-hand side of (6.20) is independent of these parameters. In terms of (6.3) we thus find that

$$\left[\frac{\partial L}{\partial t^\alpha}\xi_s^\alpha + \frac{\partial L}{\partial x^j}\zeta_s^j + \frac{\partial L}{\partial \dot{x}_\alpha^j}\left(\frac{\partial \check{x}_\alpha^j}{\partial w^s}\right)_{w^s = 0}\right](D)_{w^s = 0} + L\left(\frac{\partial D}{\partial w^s}\right)_{w^s = 0} = 0. \tag{6.21}$$

In order to exploit this conclusion it is necessary to evaluate the various partial derivatives with respect to $w^s$ which occur in (6.21).

First, if we differentiate (6.18) with respect to $w^s$, noting that according to (6.8) the variables $\dot{x}_\beta^h$ are independent of $w^s$, we obtain

$$\frac{\partial^2 \bar{x}^j}{\partial w^s \partial t^\beta} + \frac{\partial^2 \bar{x}^j}{\partial w^s \partial x^h}\dot{x}_\beta^h = \frac{\partial \check{x}_\alpha^j}{\partial w^s}\left(\frac{\partial \bar{t}^\alpha}{\partial t^\beta} + \frac{\partial \bar{t}^\alpha}{\partial x^h}\dot{x}_\beta^h\right) + \check{x}_\alpha^j\left(\frac{\partial^2 \bar{t}^\alpha}{\partial w^s \partial t^\beta} + \frac{\partial^2 \bar{t}^\alpha}{\partial w^s \partial x^h}\dot{x}_\beta^h\right).$$

In this result we now put $w^s = 0$, observing (6.5), (6.6), (6.4), and (6.19), which yields

$$\frac{\partial \zeta_s^j}{\partial t^\beta} + \frac{\partial \zeta_s^j}{\partial x^h}\dot{x}_\beta^h = \left(\frac{\partial \check{x}_\alpha^j}{\partial w^s}\right)_{w^s = 0}\delta_\beta^\alpha + \dot{x}_\alpha^j\left(\frac{\partial \xi_s^\alpha}{\partial t^\beta} + \frac{\partial \xi_s^\alpha}{\partial x^h}\dot{x}_\beta^h\right),$$

or

$$\left(\frac{\partial \check{x}_\alpha^j}{\partial w^s}\right)_{w^s = 0} = \frac{d\zeta_s^j}{dt^\alpha} - \dot{x}_\beta^j\frac{d\xi_s^\beta}{dt^\alpha}. \tag{6.22}$$

Second, denoting by $T_\alpha^\beta$ the cofactor of the element $d\bar{t}^\alpha/dt^\beta$ in the determinant (6.13), we have, in terms of (6.10):

$$\frac{\partial D}{\partial w^s} = \frac{\partial}{\partial w^s}\left(\frac{d\bar{t}^\alpha}{dt^\beta}\right)T_\alpha^\beta = \left(\frac{\partial^2 \bar{t}^\alpha}{\partial w^s \partial t^\beta} + \frac{\partial^2 \bar{t}^\alpha}{\partial w^s \partial x^h}\dot{x}_\beta^h\right)T_\alpha^\beta. \tag{6.23}$$

However, it is evident from (6.12) and (6.14) that $T_\alpha^\beta = \delta_\alpha^\beta$ when $w^s = 0$. Thus if we put $w^s = 0$ in (6.23), noting (6.6) once more, we obtain

$$\left(\frac{\partial D}{\partial w^s}\right)_{w^s = 0} = \left(\frac{\partial \xi_s^\alpha}{\partial t^\beta} + \frac{\partial \xi_s^\alpha}{\partial x^h}\dot{x}_\beta^h\right)\delta_\alpha^\beta = \frac{d\xi_s^\alpha}{dt^\alpha}. \tag{6.24}$$

We are now in a position to substitute from (6.14), (6.22), and (6.24) in (6.21). This yields

$$\frac{\partial L}{\partial t^\alpha}\xi_s^\alpha + \frac{\partial L}{\partial x^j}\zeta_s^j + \frac{\partial L}{\partial \dot{x}_\alpha^j}\left(\frac{d\zeta_s^j}{dt^\alpha} - \dot{x}_\beta^j\frac{d\xi_s^\beta}{dt^\alpha}\right) + L\frac{d\xi_s^\alpha}{dt^\alpha} = 0. \tag{6.25}$$

This is the result which we have been seeking: *it represents an identity which the Lagrangian L and its derivatives must satisfy if the fundamental integral* (5.9) *is to be invariant under the r-parameter group* (6.1). It should be observed that the quantities $\xi_s^\alpha$, $\zeta_s^j$, and their derivatives, which appear in (6.25), are prescribed by the given transformation group. We shall refer to (6.25) as an *invariance identity*.

The identity (6.25) gives rise directly to Noether's theorem in the form in which it is usually expressed. In order to derive this form we note that

$$\frac{\partial L}{\partial t^\alpha}\xi_s^\alpha = \frac{dL}{dt^\alpha}\xi_s^\alpha - \frac{\partial L}{\partial x^j}\dot{x}_\alpha^j\xi_s^\alpha - \frac{\partial L}{\partial \dot{x}_\beta^j}\ddot{x}_{\alpha\ \beta}^j\xi_s^\alpha,$$

together with

$$\frac{\partial L}{\partial \dot{x}_\alpha^j}\frac{d\zeta_s^j}{dt^\alpha} = \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}_\alpha^j}\zeta_s^j\right) - \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}_\alpha^j}\right)\zeta_s^j,$$

and

$$\frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta \frac{d\xi^\beta_s}{dt^\alpha} + \frac{\partial L}{\partial \dot{x}^j_\alpha} \ddot{x}_\alpha{}^j{}_\beta \xi^\beta_s = \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta \xi^\beta_s\right) - \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right)\dot{x}^j_\beta \xi^\beta_s.$$

These relations are substituted in turn in (6.25); on observing that

$$\frac{\partial L}{\partial \dot{x}^j_\beta} \ddot{x}_\alpha{}^j{}_\beta \xi^\alpha_s = \frac{\partial L}{\partial \dot{x}^j_\alpha} \ddot{x}_\alpha{}^j{}_\beta \xi^\beta_s,$$

it is found after a little simplification that

$$\frac{d}{dt^\alpha}\left(L\xi^\alpha_s + \frac{\partial L}{\partial \dot{x}^j_\alpha} \zeta^j_s - \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta \xi^\beta_s\right) + \left[\frac{\partial L}{\partial x^j} - \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right)\right](\zeta^j_s - \dot{x}^j_\beta \xi^\beta_s) = 0,$$

or, in terms of the Euler-Lagrange vector (5.24),

$$E_j(L)(\zeta^j_s - \dot{x}^j_\beta \xi^\beta_s) = -\frac{d\theta^\alpha_s}{dt^\alpha}, \tag{6.26}$$

where we have put

$$\theta^\alpha_s = -\left(L\xi^\alpha_s - \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta \xi^\beta_s + \frac{\partial L}{\partial \dot{x}^j_\alpha} \zeta^j_s\right). \tag{6.27}$$

The conclusion (6.26) can now be formulated as follows: *if the fundamental integral (5.9) is invariant under an r-parameter group of transformations, there exist r distinct linear combinations of the Euler-Lagrange vector $E_j(L)$, each of which is a divergence.*

The quantities $\theta^\alpha_s$ as defined by (6.27) can be expressed in a more compact form in terms of the Hamiltonian complex (5.43), namely, as

$$\theta^\alpha_s = H^\alpha_\beta \xi^\beta_s - \frac{\partial L}{\partial \dot{x}^j_\alpha} \zeta^j_s. \tag{6.28}$$

By means of (6.26) we can now formulate the *theorem of Noether* in the following form:

*If the fundamental integral (5.9) is invariant under the r-parameter transformation group (6.1), then on any extremal subspace $C_m$ the divergences of the r quantities $\theta^\alpha_s$ as defined by (6.27) or (6.28) vanish on $C_m$.*

In any physical field theory the vanishing of a divergence is usually interpreted as a conservation law; the importance of Noether's theorem is therefore due to the fact that, for any field theory based on a variational principle, it immediately suggests appropriate conservation laws. This procedure is illustrated by the following examples.

EXAMPLE 1

Let us consider the *m*-parameter transformation

$$\bar{x}^j = x^j, \qquad \bar{t}^\alpha = t^\alpha + w^\alpha. \tag{6.29}$$

According to (6.3) we have, in this case, $\zeta^j_s = 0$, $\xi^\alpha_\beta = \delta^\alpha_\beta$. The invariance identity (6.25) immediately yields

$$\frac{\partial L}{\partial t^\alpha} = 0; \tag{6.30}$$

while (6.28) reduces to

$$\theta^\alpha_\beta = H^\alpha_\beta. \tag{6.31}$$

According to Noether's theorem the divergences $dH^\alpha_\beta/dt^\alpha$ vanish identically on any extremal subspace $C_m$. This conclusion confirms the relation (5.53) when (6.30) is taken into account.

EXAMPLE 2

Into any parameter transformation of the type (5.5) one can always inject *r* parameters $w^s$ in an arbitrary manner. This gives rise to an *r*-parameter group of the type

$$\bar{t}^\alpha = \bar{t}^\alpha(t^\beta, w^s), \qquad \bar{x}^j = x^j, \tag{6.32}$$

for which $\zeta^j_s = 0$ while $\xi^\alpha_s$, $\partial\xi^\alpha_s/\partial t^\beta$ are essentially arbitrary. The invariance identity (6.25) corresponding to (6.32) reduces to

$$\frac{\partial L}{\partial t^\alpha} \xi^\alpha_s - \left(\frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta - L\delta^\alpha_\beta\right)\frac{\partial \xi^\beta_s}{\partial t^\alpha} = 0. \tag{6.33}$$

Because of the arbitrary nature of the coefficients $\xi^\alpha_s$, $\partial\xi^\beta_s/\partial t^\alpha$ this is possible if and only if

$$\frac{\partial L}{\partial t^\alpha} = 0, \qquad \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta = L\delta^\alpha_\beta.$$

Thus if the fundamental integral (5.9) is to be invariant under all parameter transformations of the type (5.5), the derivatives of *L* must satisfy these conditions. [This establishes the assertion made in the preceding section in respect of (5.15) and (5.16).]

## 6.7  HIGHER-ORDER PROBLEMS IN THE CALCULUS OF VARIATIONS

In the case of some physical applications, such as those in the general theory of relativity, the formulation of our multiple integral problem as given in Section 6.5 is still too restrictive. This is due to the fact that the integrand of the fundamental integral (5.9) depends only on the first derivatives $\dot{x}^j_\alpha$ of

the functions $x^j = x^j(t^\alpha)$ which define the subspaces $C_m$, whereas the formulation of some problems requires that the Lagrangian depend on higher-order derivatives of $x^j(t^\alpha)$. Thus this section is devoted to a very brief discussion of the so-called second-order variational problems which depend on Lagrangians of the form

$$L = L(t^\alpha, x^j, \dot{x}^j_\alpha, \ddot{x}^j_{\alpha\,\beta}), \tag{7.1}$$

whose supplementary arguments $\ddot{x}^j_{\alpha\,\beta}$ are defined by (5.8).

The formulation of such problems is identical with that given in Section 6.5, subject only to the stipulation that the boundary condition (5.11) for admissible surfaces of comparison be augmented by the condition

$$\dot{x}^j_\alpha(t^\beta) = \dot{x}^j_\alpha(t^\beta) = a^j_\alpha(t^\beta) \quad \text{for all } t^\beta \in \partial G, \tag{7.2}$$

where the functions $a^j_\alpha(t^\beta)$ are prescribed on the boundary $\partial G$ of $G$. It is assumed that the Lagrangian (7.1) is of class $C^2$ in all its arguments and that it is a scalar under coordinate transformations $\bar{x}^j = \bar{x}^j(x^h)$ on $X_n$. Under these circumstances $\partial L / \partial \ddot{x}^j_{\alpha\,\beta}$ is a type (0, 1) tensor, while $\partial L / \partial \dot{x}^j_\alpha$, $\partial L / \partial x^j$ are not tensorial.

The method of equivalent integrals may be applied as before, provided that the integrand $\Delta$ of the requisite independent integral be defined as

$$\Delta(t^\beta, x^h, \dot{x}^h_\beta, \ddot{x}^h_{\varepsilon\,\beta}) = \det(c^\alpha_\beta), \tag{7.3}$$

where

$$c^\alpha_\beta = \frac{dS^\alpha}{dt^\beta} = \frac{\partial S^\alpha}{\partial t^\beta} + \frac{\partial S^\alpha}{\partial x^h} \dot{x}^h_\beta + \frac{\partial S^\alpha}{\partial \dot{x}^h_\varepsilon} \ddot{x}^h_{\varepsilon\,\beta} \tag{7.4}$$

for some given set of $m$ class $C^2$ functions $S^\alpha = S^\alpha(t^\beta, x^h, \dot{x}^h_\beta)$. The proof of the fact that the integral of $\Delta$ is independent in the sense of (5.34) proceeds along the same lines as in Section 6.5 provided that the additional boundary conditions (7.2) are taken into account.

Again, the construction of the geodesic field entails that

$$L(t^\beta, x^h, \dot{x}^h_\beta, \ddot{x}^h_{\varepsilon\,\beta}) - \Delta(t^\beta, x^h, \dot{x}^h_\beta, \ddot{x}^h_{\varepsilon\,\beta}) = 0 \tag{7.5}$$

for such fields, while this expression is positive otherwise, and this implies that

$$\frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\beta}} = \frac{\partial \Delta}{\partial \ddot{x}^j_{\alpha\,\beta}}, \qquad \frac{\partial L}{\partial \dot{x}^j_\alpha} = \frac{\partial \Delta}{\partial \dot{x}^j_\alpha} \tag{7.6}$$

for geodesic fields. In terms of the cofactors $C^\alpha_\beta$ of the elements $c^\beta_\alpha$ of the determinant $\Delta$ these conditions can be expressed in the form

$$\frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\beta}} = C^\gamma_\varepsilon \frac{\partial c^\varepsilon_\gamma}{\partial \ddot{x}^j_{\alpha\,\beta}}, \qquad \frac{\partial L}{\partial \dot{x}^j_\alpha} = C^\gamma_\varepsilon \frac{\partial c^\varepsilon_\gamma}{\partial \dot{x}^j_\alpha}. \tag{7.7}$$

But from (7.4) we have, by differentiation with respect to $\ddot{x}^j_{\alpha\,\beta}$, observing the

symmetry of the latter in its lower indices,

$$\frac{\partial c^\varepsilon_\gamma}{\partial \ddot{x}^j_{\alpha\,\beta}} = \frac{1}{2} \frac{\partial S^\varepsilon}{\partial \dot{x}^h_\sigma} \delta^h_j (\delta^\alpha_\sigma \delta^\beta_\gamma + \delta^\alpha_\gamma \delta^\beta_\sigma) = \frac{1}{2} \left( \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \delta^\beta_\gamma + \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\beta} \delta^\alpha_\gamma \right),$$

while

$$\frac{\partial c^\varepsilon_\gamma}{\partial \dot{x}^j_\alpha} = \frac{\partial^2 S^\varepsilon}{\partial \dot{x}^j_\alpha \partial t^\gamma} + \frac{\partial^2 S^\varepsilon}{\partial \dot{x}^j_\alpha \partial x^h} \dot{x}^h_\gamma + \frac{\partial^2 S^\varepsilon}{\partial \dot{x}^j_\alpha \partial \dot{x}^h_\sigma} \ddot{x}^h_{\sigma\,\gamma} + \frac{\partial S^\varepsilon}{\partial x^j} \delta^\alpha_\gamma = \frac{d}{dt^\gamma} \left( \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) + \frac{\partial S^\varepsilon}{\partial x^j} \delta^\alpha_\gamma.$$

When these values are substituted in (7.7), we obtain

$$\frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\beta}} = \frac{1}{2} \left( C^\beta_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\beta} \right), \tag{7.8}$$

together with

$$\frac{\partial L}{\partial \dot{x}^j_\alpha} = C^\gamma_\varepsilon \frac{d}{dt^\gamma} \left( \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial x^j}. \tag{7.9}$$

However, in the same way as (5.47) is deduced from (5.46), it may be shown that

$$\frac{dC^\gamma_\varepsilon}{dt^\gamma} = 0 \tag{7.10}$$

identically, and hence (7.9) can be written as

$$\frac{\partial L}{\partial \dot{x}^j_\alpha} = \frac{d}{dt^\gamma} \left( C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial x^j}. \tag{7.11}$$

Moreover, by means of (7.4) the relation (7.3) can be expressed as

$$L\delta^\alpha_\beta = \Delta \delta^\alpha_\beta = C^\alpha_\varepsilon c^\varepsilon_\beta = C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial x^j} \dot{x}^j_\beta + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} \ddot{x}^j_{\beta\,\gamma},$$

in which we substitute from (7.11):

$$L\delta^\alpha_\beta = C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta - \frac{d}{dt^\gamma} \left( C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) \dot{x}^j_\beta + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} \ddot{x}^j_{\beta\,\gamma}$$

$$= C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta - \frac{1}{2} \frac{d}{dt^\gamma} \left( C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) \dot{x}^j_\beta$$

$$- \frac{1}{2} \frac{d}{dt^\gamma} \left( C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} \right) \dot{x}^j_\beta + \frac{1}{2} \frac{d}{dt^\gamma} \left( C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} \right) \dot{x}^j_\beta + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} \ddot{x}^j_{\beta\,\gamma}$$

$$= C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + \frac{\partial L}{\partial \dot{x}^j_\alpha} \dot{x}^j_\beta - \frac{1}{2} \frac{d}{dt^\gamma} \left( C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} \right) \dot{x}^j_\beta$$

$$+ \frac{1}{2} \left( C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} + C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) \ddot{x}^j_{\beta\,\gamma} + \frac{1}{2} \frac{d}{dt^\gamma} \left[ \left( C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} - C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} \right) \dot{x}^j_\beta \right].$$

By means of (7.8) this can be written in the form

$$L\delta^\alpha_\beta - \frac{\partial L}{\partial \dot{x}^j_\alpha}\dot{x}^j_\beta + \frac{d}{dt^\gamma}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\right)\dot{x}^j_\beta - \frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\ddot{x}_\beta{}^j{}_\gamma$$
$$= C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + \frac{1}{2}\frac{d}{dt^\gamma}\left[\left(C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} - C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha}\right)\dot{x}^j_\beta\right]. \tag{7.12}$$

This suggests that, in analogy with (5.43), we define the *Hamiltonian complex of our second-order problem by*

$$H^\alpha_\beta = -L\delta^\alpha_\beta + \frac{\partial L}{\partial \dot{x}^j_\alpha}\dot{x}^j_\beta - \frac{d}{dt^\gamma}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\right)\dot{x}^j_\beta + \frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\ddot{x}_\beta{}^j{}_\gamma, \tag{7.13}$$

while, for the sake of brevity, we shall put

$$h^{\alpha\gamma}_\beta = \frac{1}{2}\left(C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\gamma} - C^\gamma_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha}\right)\dot{x}^j_\beta. \tag{7.14}$$

Thus the final form of (7.12) is simply

$$-H^\alpha_\beta = C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial t^\beta} + \frac{d}{dt^\gamma}(h^{\alpha\gamma}_\beta). \tag{7.15}$$

The equations (7.8), (7.11), and (7.15) characterize the geodesic field of our second-order problem. In particular, it should be noted that $h^{\alpha\gamma}_\beta$ is skew-symmetric in $\alpha$ and $\gamma$ by virtue of its definition (7.14), and hence

$$\frac{d}{dt^\alpha}\left(\frac{dh^{\alpha\gamma}_\beta}{dt^\gamma}\right) = 0 \tag{7.16}$$

identically. Thus the additional term which appears on the right-hand side of (7.15) [as compared with the corresponding equation (5.44) of the first-order problem] is such that its divergence vanishes identically.

Now from (7.8) we infer that

$$\frac{d^2}{dt^\alpha\,dt^\beta}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\beta}\right) = \frac{1}{2}\frac{d^2}{dt^\alpha\,dt^\beta}\left(C^\beta_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha} + C^\alpha_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\beta}\right) = \frac{d^2}{dt^\alpha\,dt^\beta}\left(C^\beta_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha}\right),$$

while differentiation of (7.11) gives, in view of (7.10),

$$\frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^j_\alpha}\right) = \frac{d^2}{dt^\alpha\,dt^\beta}\left(C^\beta_\varepsilon \frac{\partial S^\varepsilon}{\partial \dot{x}^j_\alpha}\right) + C^\alpha_\varepsilon \frac{d}{dt^\alpha}\left(\frac{\partial S^\varepsilon}{\partial x^j}\right),$$

so that

$$\frac{d}{dt^\alpha}\left[\frac{\partial L}{\partial \dot{x}^j_\alpha} - \frac{d}{dt^\beta}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\beta}\right)\right] = C^\alpha_\varepsilon \frac{d}{dt^\alpha}\left(\frac{\partial S^\varepsilon}{\partial x^j}\right) = C^\alpha_\varepsilon \frac{\partial}{\partial x^j}\left(\frac{dS^\varepsilon}{dt^\alpha}\right) = C^\alpha_\varepsilon \frac{\partial c^\varepsilon_\alpha}{\partial x^j} = \frac{\partial \Delta}{\partial x^j}, \tag{7.17}$$

where, in the last two steps, we have used (7.4) and (1.3.15). However, because of (7.6), it follows from (7.3) that for the geodesic field

$$\frac{\partial L}{\partial x^j} = \frac{\partial \Delta}{\partial x^j}. \tag{7.18}$$

Thus (7.17) can be expressed in the form

$$E_j(L) = 0, \tag{7.19}$$

where

$$E_j(L) = \frac{d}{dt^\alpha}\left[\frac{\partial L}{\partial \dot{x}^j_\alpha} - \frac{d}{dt^\beta}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\beta}\right)\right] - \frac{\partial L}{\partial x^j}. \tag{7.20}$$

The equations (7.19) are clearly the *Euler-Lagrange equations of our second-order problem in the calculus of variations*; they represent necessary conditions which a subspace $C_m$ must satisfy in order to afford an extreme value to the fundamental integral. It should be stressed, moreover, that *in general the Euler-Lagrange equations* (7.19) *are a system of n fourth-order partial differential equations for the functions* $x^j = x^j(t^\alpha)$. Again, any subspace $C_m$ satisfying these equations is called an *extremal*.

For such subspaces there exists a direct analogue of the relation (5.53) for first-order problems, provided that the Hamiltonian complex (7.13) is used. For, if the latter is differentiated with respect to $t^\alpha$ one obtains

$$\frac{dH^\alpha_\beta}{dt^\alpha} = -\frac{dL}{dt^\beta} + \frac{d}{dt^\alpha}\left[\frac{\partial L}{\partial \dot{x}^j_\alpha} - \frac{d}{dt^\gamma}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\right)\right]\dot{x}^j_\beta + \frac{\partial L}{\partial \dot{x}^j_\alpha}\ddot{x}_\alpha{}^j{}_\beta$$
$$- \frac{d}{dt^\gamma}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\right)\ddot{x}_\alpha{}^j{}_\beta + \frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\right)\ddot{x}_\beta{}^j{}_\gamma + \frac{\partial L}{\partial \ddot{x}_\alpha{}^j{}_\gamma}\frac{d\ddot{x}_{\beta\gamma}}{dt^\alpha}$$

from which it follows by simplification after expansion of $dL/dt^\beta$ and use of (7.20) that

$$\frac{dH^\alpha_\beta}{dt^\alpha} = -\frac{\partial L}{\partial t^\beta} + E_j(L)\dot{x}^j_\beta. \tag{7.21}$$

This is an identity: for an extremal subspace, however,

$$\frac{dH^\alpha_\beta}{dt^\alpha} = -\frac{\partial L}{\partial t^\beta}. \tag{7.22}$$

Incidentally, the latter relation may also be deduced directly from (7.15) by virtue of (7.16), (7.10), and (7.5).

## PROBLEMS

6.1   Let $n = 1$, and suppose that the Lagrangian is defined to be $L = t^{-1}\sqrt{1 + \dot{x}^2}$. Show that the extremal passing through the points $(t = 1, x = 0)$ and $(t = 2, x = 1)$ is the circle $(x - 2)^2 + t^2 = 5$.

6.2   Why does the Lagrangian $L = x^2 + 2tx\dot{x}$ fail to give rise to extremal curves?

6.3   If $L = F(x^j, \dot{x}^j)$, that is, $L$ is independent of $t$, show that the Euler-Lagrange equations (2.20) always possess a first integral

$$F - \dot{x}^j \frac{\partial F}{\partial \dot{x}^j} = C.$$

6.4   Show that a first integral of the Euler-Lagrange equations associated with the Lagrangian $L = x^2 + \sqrt{1 + \dot{x}^2}$ is given by

$$(A - x^2)\sqrt{1 + \dot{x}^2} = 1,$$

where $A$ is a constant.

6.5   Let $n = 1$, and let $C$ denote a curve joining two given points $P_1(t_1, x_1)$, $P_2(t_2, x_2)$. Prove the following assertion: in order that the area of the surface of revolution obtained by rotating $C$ about the $t$-axis assumes a minimum value relative to all other surfaces thus obtained by means of the family of curves joining $P_1$ and $P_2$, it is necessary that $C$ be a catenary: $x = a \cosh[(t + b)a^{-1}]$, where $a$ and $b$ are constants.

6.6   Let $n = 1$, and show that the Hamiltonian function associated with the Lagrangian $L(t, x, \dot{x}) = f(t, x) (1 + \dot{x}^2)^{1/2}$ is given by $H = -\sqrt{f^2 - p^2}$. Write down the canonical equations, and show that if, in particular, $f^2 = x^2 + \phi(t)$, where $\phi(t)$ is arbitrary, then the canonical equations possess a first integral of the form $p^2 - x^2 = a$, where $a$ is a constant. (It may be assumed that $f(t, x) > 0$.)

6.7   Show that the Hamilton-Jacobi equation associated with the Lagrangian of Problem 6.6 is equivalent to

$$\left(\frac{\partial S}{\partial t}\right)^2 + \left(\frac{\partial S}{\partial x}\right)^2 = f^2(t, x).$$

If $\partial f/\partial t = 0$, a solution is given by

$$S(t, x) = At + \int_{x_0}^{x} \sqrt{f^2(\tau) - A^2} \, d\tau + B$$

where $A$, $B$ are constants. Deduce that, for the extremals, $\dot{x} = A^{-1}\sqrt{f^2 - A^2}$.

6.8   (a) By examining the behavior of the excess function associated with the Lagrangian (2.49) of the harmonic oscillator, establish whether the corresponding variational principle is concerned with maxima or minima of the fundamental integral.
    (b) Show that the family of transversals does not, at any point, intersect the family of the extremals of the harmonic oscillator orthogonally.

6.9   Show that, corresponding to $L = \sqrt{t^2 + x^2}\sqrt{1 + \dot{x}^2}$, the extremals for which $x = A$, $dx/dt = 0$ when $t = 0$ are given by $x^2 - t^2 = A^2$. Show that the family of curves which are transversal to the extremal field is characterized by $xt = B$.

6.10   Let $(t, x^j)$ denote rectangular coordinates in $E_3$ $(n = 1, 2)$, and let $\mu(t, x^j)$ represent the refractive index of an inhomogeneous isotropic optical medium. Interpret the Lagrangian $L = c^{-1}\mu(t, x^i) [1 + (\dot{x}^1)^2 + (\dot{x}^2)^2]^{1/2}$ in terms of Fermat's principle (where $c$ denotes the velocity of light in vacuo) and write down the corresponding Euler-Lagrange equation.

6.11   If $L(t, x^i, \dot{x}^i) = \mathscr{L}(t, x^i, \dot{x}^i) + A_i \dot{x}^i$ where $A_i = A_i(x^h)$ and

$$\frac{\partial A_i}{\partial x^j} - \frac{\partial A_j}{\partial x^i} = 0,$$

show that

$$E_i(L) = E_i(\mathscr{L}).$$

Explain why $A_i \dot{x}^i$ does not contribute to the Euler-Lagrange expression.

*6.12   If $L = L(t, x^i, \dot{x}^i)$ is such that $E_i(L)$ vanishes identically show that there exists an $\alpha = \alpha(t, x^h)$ for which $L = d\alpha/dt$.

6.13   Show that if $L = L(t, x^j, \dot{x}^j)$ and

$$E_i(L) = f\dot{x}^i + g^i$$

where $f = f(t, x^j)$ and $g^i = g^i(t, x^j)$ then $f \equiv 0$.

6.14   A particle of mass $m$ in a three-dimensional Euclidean space $E_3$ referred to Cartesian coordinates $x^1, x^2, x^3$ is subjected to a central force represented by a potential function $V(t, r)$, where $r^2 = (x^1)^2 + (x^2)^2 + (x^3)^2$. The Lagrangian of this system is given by $L = \frac{1}{2}m\dot{x}^i\dot{x}^i - V(t, r)$. Show that the corresponding fundamental integral is invariant under the three-parameter group of transformations (where $\omega^l$ are constants and terms of order $\omega^l\omega^h$ are neglected)

$$\bar{x}^j = x^j + \varepsilon_{jhl}x^h \omega^l.$$

It should be noted that this represents an infinitesimal orthogonal transformation or rotation [see condition (2.1.5)]. By means of Noether's theorem deduce that this implies the conservation of the angular momentum $\boldsymbol{\theta} = \mathbf{r} \times \mathbf{p}$ of the particle about the origin.

6.15   Let $L(t, x^h, \dot{x}^h) = g(t, x^h)l(\dot{x}^h)$ where

$$g(t, x^h) > 0 \quad \text{and} \quad l(\dot{x}^h) = \left[1 + \sum_{j=1}^{n} \dot{x}^j\dot{x}^j\right]^{1/2}.$$

(a) Show that

$$E(t, x^h, \psi^h, \dot{x}^h) = gl(\dot{x}^h)\left[1 - \frac{1 + \sum_{h=1}^{n} \psi^h\dot{x}^h}{l(\psi^j)l(\dot{x}^k)}\right]$$

and deduce that $E > 0$ for all $(t, x^h, \dot{x}^h)$ with $\psi^h \neq \dot{x}^h$.

(b) Show that

$$\frac{d}{dt}\left(\frac{g}{l}\right) = l\frac{\partial g}{\partial t}$$

along an extremal.

6.16 Show that (5.25) is equivalent to

$$(m-1)!\,\frac{\partial L}{\partial x^i}\,dt^1 \wedge \cdots \wedge dt^m = d\!\left[\varepsilon_{\alpha\alpha_2\cdots\alpha_m}\frac{\partial L}{\partial \dot{x}^i_\alpha}\,dt^{\alpha_2} \wedge \cdots \wedge dt^{\alpha_m}\right].$$

6.17 Find the Euler-Lagrange equations associated with the Lagrangians ($n = 1$, $m = 2$)

(a) $L = \rho(\dot{x}^1)^2 - \tau(\dot{x}^2)^2$ ($\rho, \tau$ constants; vibrating string problem);

(b) $L = [1 + (\dot{x}^1)^2 + (\dot{x}^2)^2]^{1/2}$ (minimal surface problem in $E_3$);

(c) $L = (\dot{x}^1)^2 + (\dot{x}^2)^2 + 2xf(t^1, t^2)$.

6.18 If $L(t^\alpha, x^j, \dot{x}^j_\alpha) = \sum_{\alpha=1}^{m}(d\lambda^\alpha/dt^\alpha)$ where $\lambda^\alpha = \lambda^\alpha(t^\beta, x^j)$, show that $E_j(L)$, defined by (5.24), vanishes identically. Generalize this result to $L(t^\alpha, x^j, \dot{x}^j_\alpha, \ddot{x}^j_{\alpha\beta})$.

6.19 Find the Euler–Lagrange equation associated with

(a) $L(t^\alpha, x, \dot{x}_\alpha) = \sum_{\alpha=1}^{n}\dot{x}_\alpha\dot{x}_\alpha,$

and

(b) $L(t^\alpha, x, \dot{x}_\alpha, \ddot{x}_{\alpha\beta}) = -\sum_{\alpha=1}^{n}x\ddot{x}_{\alpha\alpha}.$

Comment on these equations.

6.20 Show by direct transformation that, if the Lagrangian (7.1) is a scalar under coordinate transformations $\bar{x}^j = \bar{x}^j(x^h)$ on $X_n$, then

$$\frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\beta}}, \qquad \frac{\partial L}{\partial \dot{x}^j_\alpha} - 2\frac{d}{dt^\beta}\!\left(\frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\beta}}\right), \qquad E_j(L),$$

the latter being defined according to (7.20), are each the components of type (0, 1) tensors.

6.21 Show that if the integral of the Lagrangian (7.1) is to be parameter-invariant (that is, invariant under arbitrary transformations $\bar{t}^\alpha = \bar{t}^\alpha(t^\varepsilon)$ with nonvanishing determinant), then $L$ must satisfy the conditions

$$\frac{\partial L}{\partial t^\alpha} = 0, \qquad \frac{\partial L}{\partial \dot{x}^j_\alpha}\dot{x}^j_\beta + 2\frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\varepsilon}}\ddot{x}^j_{\beta\,\varepsilon} = \delta^\alpha_\beta L, \qquad \frac{\partial L}{\partial \ddot{x}^j_{\alpha\,\beta}}\dot{x}^j_\varepsilon = 0.$$

# 7
# RIEMANNIAN GEOMETRY

The theory of tensors and differential forms as developed in Chapters 3–5 is based on the assumption that our differentiable manifold $X_n$ is endowed solely with an affine connection. Metric concepts, such as the length of a displacement, or the $m$-dimensional volume element ($m \le n$), do not occur within the context of the general theory. In this respect the above treatment differs radically from the historical development of the subject, for originally the tensor calculus arose from the theory of Riemannian spaces. The latter represent natural generalizations to $n$ dimensions of smooth two-dimensional surfaces embedded in an $E_3$, these surfaces being differentiable manifolds endowed with a metric in the sense that there exists a type (0, 2) symmetric tensor field $g_{hj}$ in terms of which the lengths of tangent vectors are defined. (In the classical theory of surfaces the components $g_{11}$, $g_{12}$, $g_{21}$, $g_{22}$ are usually denoted by $E$, $F$, $F$, $G$, respectively, and the quadratic form defined by these coefficients is called the first fundamental form of the surface.) Similarly, an $n$-dimensional Riemannian space $V_n$ is a differentiable manifold which is endowed, in the first instance, with a nonsingular type (0, 2) symmetric tensor field, the so-called metric tensor field. By means of this metric the arc length of a differentiable curve on $V_n$ is defined as a single integral, which gives rise in a natural manner to a parameter-invariant variational problem. The extremals of the latter are called the geodesics of $V_n$, and an analysis of the corresponding Euler-Lagrange equation reveals the existence of a connection, whose coefficients, the so-called Christoffel symbols, are uniquely expressible in terms of the partial derivatives of the metric tensor field. Thus, on the one hand, the existence of a metric ensures the existence of an associated connection together with a corresponding theory of curvature, whereas on the other hand, the existence of a connection alone does not

in general imply the existence of a metric. The former state of affairs reflects the historical development of the tensor calculus; in fact, it was realized only much later that the mere existence of a connection is sufficient for the purposes of a self-contained theory of tensors.

In the preceding sections the latter approach is adopted, largely because it is our aim to present the tensor calculus as a discipline which is essentially independent of the background of metric differential geometry. This, however, has entailed certain painful losses, for the presence of a metric guarantees a much richer geometrical structure than that which is implied by a connection alone, and this in turn allows for a more appealing and concrete geometrical interpretation. Moreover, it should also be observed that the general theory of relativity is based on the assumption that space-time is a four-dimensional Riemannian space, and thus the metric theory is quite indispensable to the relativist.

In the classical theory of surfaces the quadratic form defined by the components of the metric tensor is positive definite; this assumption, however, will not be made in our treatment of Riemannian geometry. Although this relaxation gives rise to some technical complications, it is imperitive from the point of view of relativistic applications that the properties of indefinite metrics be analyzed, particularly as regards the behavior of certain types of geodesics. The theory of curvature of Riemannian spaces is of course a special case of the theory treated in Chapter 3, but it will be seen that it displays some important peculiarities which are a direct consequence of the existence of a metric. The theory of hypersurfaces of a Riemannian space is very similar to the classical theory of surfaces embedded in a three-dimensional Euclidean space, and it is hoped that the presentation given below reflects some of the beauty inherent in classical differential geometry.

## 7.1  INTRODUCTION OF A METRIC

We now endeavor to introduce the notion of a metric into our general theory (see, e.g., E. Cartan [4], Eisenhart [1], Levi-Civita [1], Schouten [1], and Synge and Schild [1]). More precisely, we shall associate a concept of length with co- and contravariant vectors, the latter being regarded once more as elements of appropriate tangent spaces of our differentiable manifold $X_n$.

Let us suppose, then, that we are given a contravariant vector in the tangent space $T_n(P)$ of some point $P$ of $X_n$. Relative to some coordinate system on $X_n$, for which the coordinates of $P$ are $x^j$, the components of this vector are denoted by $X^j$. By analogy with the expression for the length of a vector in a Euclidean space referred to a curvilinear coordinate system we shall assume that the length $|X|$ of our given vector is the modulus of some function $F(x^j, X^j)$ of the positional coordinates $x^j$ as well as of the components $X^j$.

Clearly this function must satisfy certain conditions in order to yield a satisfactory concept of length; we begin by listing the most immediate stipulations.

1. The function $F(x^j, X^j)$ is of class $C^r$ in all its arguments, where $r \geq 5$.
2. The function $F(x^j, X^j)$ is an invariant under arbitrary coordinate transformations on $X_n$.
3. The function $F(x^j, X^j)$ is positively homogeneous of the first degree in the arguments $X^j$; that is,

$$F(x^j, \lambda X^j) = \lambda F(x^j, X^j) \quad \text{for all } \lambda > 0. \tag{1.1}$$

The motivation of these three conditions is immediate; the second requires that the length of any vector be independent of the choice of the coordinate system on $X_n$, while the third ensures that the length $|\lambda X|$ of the vector with components $\lambda X^j$, for some positive number $\lambda$, be $\lambda |X| = \lambda |F(x^j, X^j)|$. Additional requirements on $F$ will be imposed presently.

This construction gives rise immediately to the concept of arc length of a differentiable curve $C$ of $X_n$. Assuming that $C$ can be represented in the form $x^j = x^j(t)$, we associate the line-element $ds = F(x^j, dx^j)$ with the contravariant vector $dx^j = \dot{x}^j \, dt$, where $\dot{x}^j = dx^j/dt$. (Here it should be noted that $ds$ is the length of the displacement defined by $dx^j$ whenever $F(x^j, dx^j) > 0$). The arc length of $C$ between two points $P_1, P_2$ corresponding to parameter values $t_1$ and $t_2$ is now defined as

$$s = \int_{C P_1}^{P_2} F(x^j, dx^j) = \int_{C t_1}^{t_2} F(x^j, \dot{x}^j) \, dt, \tag{1.2}$$

where, in the second step, we have used (1.1). In passing we note that this integral is parameter-invariant in the sense of Section 6.1.

Before proceeding with the geometrical theory, we derive a few simple analytical consequences of (1.1). From Euler's theorem on homogeneous functions we have that

$$\frac{\partial F(x, \dot{x})}{\partial \dot{x}^j} \dot{x}^j = F(x, \dot{x}), \tag{1.3}$$

$$\frac{\partial^2 F(x, \dot{x})}{\partial \dot{x}^j \, \partial \dot{x}^h} \dot{x}^h = 0, \tag{1.4}$$

so that

$$\det\left(\frac{\partial^2 F(x, \dot{x})}{\partial \dot{x}^j \, \partial \dot{x}^h}\right) = 0. \tag{1.5}$$

[Here, and in the sequel, we do not affix indices to variables which appear merely as arguments: thus $F(x, \dot{x})$ is intended to represent $F(x^j, \dot{x}^j)$.]

Furthermore, since

$$\frac{1}{2}\frac{\partial F^2(x,\dot{x})}{\partial \dot{x}^j} = F(x,\dot{x})\frac{\partial F(x,\dot{x})}{\partial \dot{x}^j},$$

it follows that

$$\frac{1}{2}\frac{\partial^2 F^2(x,\dot{x})}{\partial \dot{x}^j\,\partial \dot{x}^h} = \frac{\partial F(x,\dot{x})}{\partial \dot{x}^j}\frac{\partial F(x,\dot{x})}{\partial \dot{x}^h} + F(x,\dot{x})\frac{\partial^2 F(x,\dot{x})}{\partial \dot{x}^j\,\partial \dot{x}^h}, \tag{1.6}$$

and hence, by virtue of (1.3) and (1.4),

$$F^2(x,\dot{x}) = \frac{1}{2}\frac{\partial^2 F^2(x,\dot{x})}{\partial \dot{x}^j\,\partial \dot{x}^h}\dot{x}^j\dot{x}^h. \tag{1.7}$$

This identity suggests that we introduce the notation

$$g_{jh}(x,\dot{x}) = \frac{1}{2}\frac{\partial^2 F^2(x,\dot{x})}{\partial \dot{x}^j\,\partial \dot{x}^h}, \tag{1.8}$$

so that (1.7) can be written in the form

$$F^2(x,\dot{x}) = g_{jh}(x,\dot{x})\dot{x}^j\dot{x}^h. \tag{1.9}$$

The length $|X|$ of any contravariant vector $X^j$ of $T_n(P)$ can thus be regarded as the positive square root of

$$|X|^2 = |g_{jh}(x,X)X^jX^h|, \tag{1.10}$$

while the line-element $ds$ associated with a displacement $dx^j$ at $P$ is given by

$$ds^2 = g_{jh}(x,dx)\,dx^h\,dx^j. \tag{1.11}$$

It is easily verified that the quantities defined by (1.8) represent the components of a type (0, 2) tensor. Under a nonsingular coordinate transformation $\bar{x}^j = \bar{x}^j(x^h)$ on $X_n$, we have $\dot{x}^h = (\partial x^h/\partial \bar{x}^l)\dot{\bar{x}}^l$, so that

$$\frac{\partial \dot{x}^h}{\partial \dot{\bar{x}}^j} = \frac{\partial x^h}{\partial \bar{x}^l}\delta_j^l = \frac{\partial x^h}{\partial \bar{x}^j}, \tag{1.12}$$

while, according to condition 2,

$$\bar{F}^2(\bar{x}^j,\dot{\bar{x}}^j) = F^2(x^h,\dot{x}^h). \tag{1.13}$$

Thus, differentiating (1.13) successively with respect to $\dot{\bar{x}}^j$ and $\dot{\bar{x}}^l$, we find that

$$\frac{\partial \bar{F}^2}{\partial \dot{\bar{x}}^j} = \frac{\partial F^2}{\partial \dot{x}^h}\frac{\partial \dot{x}^h}{\partial \dot{\bar{x}}^j} = \frac{\partial F^2}{\partial \dot{x}^h}\frac{\partial x^h}{\partial \bar{x}^j},$$

$$\frac{\partial^2 \bar{F}^2}{\partial \dot{\bar{x}}^j\,\partial \dot{\bar{x}}^l} = \frac{\partial^2 F^2}{\partial \dot{x}^h\,\partial \dot{x}^k}\frac{\partial x^h}{\partial \bar{x}^j}\frac{\partial x^k}{\partial \bar{x}^l},$$

or in terms of the notation (1.8),

$$\bar{g}_{jl} = g_{hk}\frac{\partial x^h}{\partial \bar{x}^j}\frac{\partial x^k}{\partial \bar{x}^l}, \tag{1.14}$$

which is the transformation law of a type (0, 2) tensor. We note that, by definition, the tensor (1.8) is symmetric.

When various additional restrictions are imposed upon the function $F(x,\dot{x})$ one obtains certain well-known geometries. For instance, if it is required that $F(x,\dot{x}) \geq 0$, with equality if and only if $\dot{x}^1 = \dot{x}^2 = \cdots = \dot{x}^n = 0$, and that the rank of the matrix $(\partial^2 F/\partial \dot{x}^j\,\partial \dot{x}^h)$ be $n-1$, the resulting geometry is called a *Finsler geometry*, or equivalently, the manifold $X_n$ is said to be a Finsler space (see, e.g., Rund [1]).

On the other hand, if it is assumed that the tensor (1.8) depends solely on the positional coordinates $x^j$, the determinant $\det(g_{jh})$ being nonzero, one says that the manifold $X_n$ is endowed with a *Riemannian metric*, the tensor (1.8) being called the *metric tensor*. Under these circumstances the identity (1.9) reduces to

$$F^2(x,\dot{x}) = g_{jh}(x)\dot{x}^j\dot{x}^h, \tag{1.15}$$

which indicates clearly that $F^2$ is a quadratic form in the $\dot{x}^j$. Furthermore, according to (1.10), the square of the length $|X|$ of a contravariant vector is a quadratic form in the components $X^j$ of that vector, the coefficients of which depend on the coordinates of the point $P$ at which the vector is defined. [A special case of this state of affairs is given by the expression (1.2.16) for the length of a vector in a Euclidean space referred to curvilinear coordinates.]

We shall henceforth restrict our attention to *Riemannian spaces*; in the light of the above remarks these can be briefly described as differentiable manifolds endowed with a type (0, 2) tensor field whose components are denoted by $g_{jh}(x)$, the latter being symmetric and such that

$$g = \det(g_{jh}) \neq 0. \tag{1.16}$$

In passing we note that, according to the theory of Section 4.1, this determinant is a scalar density of weight 2, and consequently the condition (1.16) is independent of the choice of the coordinates. The manifold $X_n$ thus endowed with a Riemannian metric will be denoted by $V_n$. From (1.11) we infer that with each displacement $dx^j$ at a point $P$ of $V_n$ we may associate the so-called *line-element $ds$* according to

$$ds^2 = g_{jh}(x)\,dx^j\,dx^h. \tag{1.17}$$

*Remark.* The definition of a Riemannian metric as given in some texts includes the additional condition that the quadratic form (1.17) be positive definite, the manifold being called pseudo-Riemannian otherwise; however,

since many physical applications involve line-elements of the type (1.17) which are not positive definite, this restriction will not be imposed here. This implies that we have to take into account the existence of the so-called *null vectors*, which are such that their lengths as defined by (1.10) vanish, although their components are not all zero simultaneously. (The geometrical significance of such null vectors will soon become evident.)

We now investigate some simple geometrical consequences of the fact that the length $|X|$ of a contravariant vector with components $X^j$ at $P$ is given by the positive square root of

$$|X|^2 = |g_{jh}(x)X^jX^h|. \tag{1.18}$$

Let $Y^j$ denote the components of some other contravariant vector in $T_n(P)$; the sums $X^j + Y^j$ then represent another element of $T_n(P)$ for which

$$|X + Y|^2 = |g_{jh}(x)(X^j + Y^j)(X^h + Y^h)|. \tag{1.19}$$

Furthermore, the bilinear form

$$(X \cdot Y) = g_{jh}(x)X^jY^h \tag{1.20}$$

is obviously an invariant since $g_{jh}(x)$ is a type (0, 2) tensor: this quantity will be regarded as an *inner* (or scalar) product of the vectors $X^j$, $Y^j$. (We note in passing that this construction is possible only by virtue of the fact that we are in possession of a metric; on an arbitrary differentiable manifold devoid of a metric one can only define the inner product of vectors which are respectively co- and contravariant.)

In terms of (1.20) we can write

$$g_{hj}(X^h + Y^h)(X^j + Y^j) = g_{hj}X^hX^j + 2(X \cdot Y) + g_{hj}Y^hY^j. \tag{1.21}$$

Furthermore, for any real number $\lambda$ we have similarly that

$$g_{hj}(X^h + \lambda Y^h)(X^j + \lambda Y^j) = g_{hj}X^hX^j + 2\lambda(X \cdot Y) + \lambda^2 g_{hj}Y^hY^j, \tag{1.22}$$

and it now becomes necessary to distinguish between two distinct cases.

CASE 1   The quadratic form (1.17) is positive definite.

Under these circumstances the expression (1.22), regarded as a quadratic in $\lambda$, is positive for all values of $\lambda$, which is possible if and only if its discriminant is nonpositive, that is, if

$$(X \cdot Y)^2 \leq (g_{hj}X^hX^j)(g_{kl}Y^kY^l),$$

or, since $g_{hj}X^hX^j$, $g_{kl}Y^kY^l$ are positive,

$$(X \cdot Y) \leq |X||Y|. \tag{1.23}$$

It then follows from (1.21) that

$$|X + Y|^2 \leq |X|^2 + 2|X||Y| + |Y|^2,$$

or, since all entries in this inequality represent positive real numbers,

$$|X + Y| \leq |X| + |Y|, \tag{1.24}$$

which is the triangle inequality. Thus the latter, which holds for all pairs of vectors in $T_n(P)$, is a direct consequence of the assumption that the metric is positive definite.

CASE 2   The quadratic form (1.17) assumes both positive and negative values.

Let us assume that the vectors $X^j$, $Y^j$, which appear in (1.21), span a plane which contains null directions. It then follows that the quadratic (1.22) in $\lambda$ may change sign, which implies that its discriminant is nonnegative, that is,

$$(X \cdot Y)^2 \geq |X|^2|Y|^2. \tag{1.25}$$

For the sake of simplicity, let us choose two vectors for which $g_{hj}X^hX^j = 1$, $g_{hj}Y^hY^j = 1$, and for which $(X \cdot Y)$ is positive (which is always possible by continuity if $X^h$ and $Y^h$ differ by small amounts). The inequality (1.25) then reduces to $(X \cdot Y) \geq 1$. Accordingly it follows from (1.21) that $g_{hj}(X^h + Y^h)(X^j + Y^j) \geq 4$, so that $|X + Y|^2 \geq 4$, while $|X| + |Y| = 2$. Thus, for this given pair of vectors,

$$|X + Y| \geq |X| + |Y|, \tag{1.26}$$

from which it is inferred that, for the case of an indefinite metric, there exist vectors satisfying the *reversed triangle inequality*. [A similar approach can also be applied to a pair of vectors for which $g_{hj}X^hX^j = -1$, $g_{hj}Y^hY^j = -1$, and which are sufficiently close to each other to ensure that $(X \cdot Y)$ is negative; for under these circumstances (1.25) yields $(X \cdot Y) \leq -1$, and hence $g_{hj}(X^h + Y^h)(X^j + Y^j) \leq -4$, so that again $|X + Y|^2 \geq 4$, with $|X| + |Y| = 2$. However, it should be noted that the inequality (1.26) does not hold for *any* pair of vectors.]

These conclusions indicate quite clearly that the distinction between positive definite and indefinite metrics has some profound geometrical implications. These will now be investigated a little more closely for each case.

CASE 1   Positive definite case.

A contravariant vector of $T_n(P)$ for which $|X| = 1$ is said to be of *unit length*. From (1.18) it follows that all unit vectors in $T_n(P)$ satisfy the condition

$$g_{jh}(x)X^jX^h = 1. \tag{1.27}$$

Regarded as a locus in $T_n(P)$—for which the coordinates $x^j$ of $P$ are fixed, so that the $g_{jh}$ may be regarded as constant coefficients, while the $X^j$ are variable—the equation (1.27) represents an $(n-1)$-dimensional ellipsoid, or hyperellipsoid, which is obviously a generalization of the concept of *unit sphere*. The fact that the locus (1.27) is an ellipsoid and no other quadric hypersurface is due to our assumption that the left-hand side of (1.27) is positive definite.

Let $X$, $Y$ be two unit vectors, represented as directed segments in $T_n(P)$ as indicated in Figure 8. The sum $X + Y$ of these vectors is identified as usual with the diagonal of the parallelogram defined by $X$, $Y$. From the convexity of the ellipsoid (which implies that all interior points of the line segment joining the end-points of the vectors $X$, $Y$ lie within the ellipsoid) it follows that the intersection of the two diagonals of the parallelogram lies within the ellipsoid, and hence that $\frac{1}{2}|X + Y| \leq 1$, with equality if and only if $X$ and $Y$ coincide. Since $|X| = 1$, $|Y| = 1$ by construction, this conclusion immediately verifies the triangle inequality (1.24) in respect of these vectors.
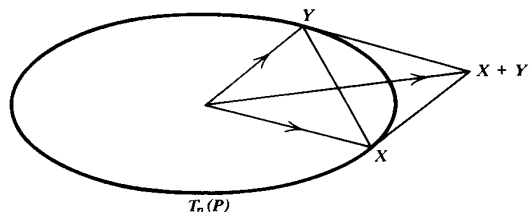


**Fig. 8**

CASE 2    Indefinite case.

Since the quadratic form $g_{jh}X^jX^h$ may assume both positive and negative values, it is convenient, for the purposes of this discussion, to introduce the so-called *indicator* $\varepsilon(X)$, which is defined by

$$\begin{aligned}\varepsilon(X) &= 1 \qquad \text{whenever } g_{jh}X^jX^h > 0,\\ \varepsilon(X) &= -1 \quad \text{whenever } g_{jh}X^jX^h < 0.\end{aligned} \tag{1.28}$$

We can now write the definition (1.18) in the form

$$|X|^2 = \varepsilon(X)g_{jh}X^jX^h, \tag{1.29}$$

recalling that $|X|$ is the positive square root of $|X|^2$. Thus $X$ is a unit vector whenever $g_{jh}X^jX^h = \pm 1$, and the counterpart of the generalized unit sphere (1.27) consists of the *pair* of hypersurfaces

$$g_{jh}X^jX^h = 1, \qquad g_{jh}X^jX^h = -1. \tag{1.30}$$

The nature of these surfaces obviously depends on the signature of the quadratic form $g_{jh}X^jX^h$. Since a general discussion of all relevant possibilities is entirely beyond our scope, we shall merely illustrate the reversed triangle inequality for the case $n = 2$ for a metric of the type

$$g_{jh}(x) = \begin{pmatrix} a_1(x) & 0 \\ 0 & -a_2(x) \end{pmatrix}, \tag{1.31}$$

it being assumed that $a_1(x) > 0$, $a_2(x) > 0$. Under these circumstances the loci (1.30) are represented by pairs of conjugate hyperbolae as illustrated in Figure 9. The asymptotes correspond to directions for which $g_{jh}X^jX^h = 0$; they therefore represent the directions of null vectors. These divide the $(X^1, X^2)$-plane into four regions, in two of which $\varepsilon(X) = 1$, while in the other two $\varepsilon(X) = -1$, as indicated. Now if $X$, $Y$ are two unit vectors both lying in one and the same region for which $\varepsilon(X) = \varepsilon(Y) = 1$ (as in Figure 9), it is evident from the concavity of the hyperbola that $\frac{1}{2}|X + Y| \geq 1$, with equality if and only if the vectors $X$, $Y$ coincide. Since $|X| = |Y| = 1$ by
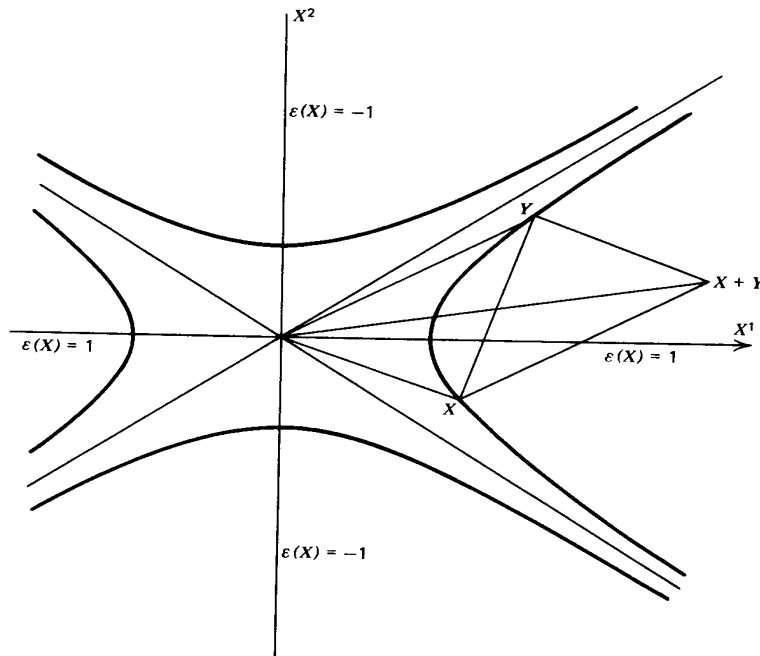


**Fig. 9**

construction, it follows that they satisfy the reversed triangle inequality (1.26). A similar conclusion is reached if $X$, $Y$ are unit vectors lying in one and the same region for which $\varepsilon(X) = \varepsilon(Y) = -1$. However, if $\varepsilon(X) = 1$, while $\varepsilon(Y) = -1$, this line of reasoning breaks down, and it is easily seen by examining special cases that any one of the three possibilities

$$|X + Y| \gtreqqless |X| + |Y|$$

is feasible. In practice, however, one is generally interested only in sums of vectors whose indicators have the same values.

The above observations should serve to emphasize the difference between geometries based respectively on positive definite and indefinite metrics. In this connection attention should be paid to a related feature. When the quadratic form (1.18) is positive definite, it is always possible to find a linear transformation on $T_n(P)$ which reduces (1.18) to a sum of squares (Hadley [1]), namely,

$$|X|^2 = (\overline{X}^1)^2 + (\overline{X}^2)^2 + \cdots + (\overline{X}^n)^2, \tag{1.32}$$

where $\overline{X}^j$ denotes the components of the vector $X$ in the new coordinate system in $T_n(P)$. This procedure enables one to regard the hyperellipsoid (1.27) as the Euclidean unit hypersphere. Accordingly the geometry of the tangent space $T_n(P)$ is essentially Euclidean. However, it should be stressed that the linear transformation which reduces the coefficients $g_{jh}(x)$ of the quadratic form (1.18) in $T_n(P)$ to Kronecker deltas $\delta_{jh}$ will not, in general, have the same effect on the coefficients $g_{jh}(x + dx)$ corresponding to the tangent space $T_n(Q)$ of a neighboring point $Q$ with coordinates $x^j + dx^j$. Accordingly the application of such affine transformations is of little practical value. Nevertheless, having realized that each tangent space $T_n(P)$ is essentially Euclidean, we shall say that *a Riemannian space $V_n$ with positive definite metric is locally Euclidean, it being understood that this metric varies from point to point of $V_n$ by virtue of the dependence of the metric tensor on its positional arguments.*

If it is possible to find a coordinate transformation in $V_n$ which is such that the transformed metric tensor $g_{jh}$ is independent of its positional coordinates, that is, such that in the transformed system $\partial \bar{g}_{jh}/\partial \bar{x}^k = 0$ everywhere, then one may, of course, introduce a linear transformation which reduces the quadratic form (1.18) to the form (1.32) in each tangent space. Under these circumstances our Riemannian space $V_n$ is simply a Euclidean space $E_n$. The latter spaces are therefore characterized among the Riemannian spaces as manifolds which admit coordinate systems in which the components of the metric tensor are everywhere constant.

More generally, allowing also for the case of indefinite metrics, we call

a space endowed with a Riemannian metric a *flat* space whenever it admits coordinate systems for which the components of the metric tensor are constants. Necessary and sufficient conditions that a space be flat in this sense are derived presently.

Thus far we have been concerned merely with the notion of length of a contravariant vector. We now indicate how this leads directly to a similar concept for covariant vectors. Let $X^h$ denote the components of an arbitrary element of $T_n(P)$. Since $g_{jh}$ is a type $(0, 2)$ tensor, it follows that the quantities defined by

$$Z_j = g_{jh}(x)X^h \tag{1.33}$$

represent the components of a covariant vector, that is, an element of the dual tangent space $T_n^*(P)$. By virtue of the hypothesis (1.16) the matrix $(g_{jh})$ possesses an inverse with entries $g^{hj}$, which are the components of a symmetric type $(2, 0)$ tensor, satisfying the conditions

$$g_{jh}(x)g^{jk}(x) = \delta_h^k. \tag{1.34}$$

Thus if we multiply (1.33) by $g^{jk}$, summing over $j$, we obtain

$$X^k = g^{jk}(x)Z_j, \tag{1.35}$$

which is obviously the inverse of (1.33). *The relations* (1.33) *and* (1.35) *establish a one-to-one correspondence between the tangent spaces $T_n(P)$ and $T_n^*(P)$.* It should be emphasized that this correspondence depends on the existence of the given metric: for an arbitrary differentiable manifold $X_n$ devoid of a metric tensor no such relationship between the tangent space and its dual exists.

Let us now substitute from (1.35) in the quadratic form $g_{lk}X^lX^k$. This gives, in view of (1.34),

$$g_{lk}X^lX^k = g_{lk}g^{lh}Z_h g^{jk}Z_j = \delta_k^h Z_h g^{jk}Z_j = g^{jh}Z_j Z_h, \tag{1.36}$$

which clearly suggests that we should define *the length $|Z|$ of any covariant vector $Z_j$* by

$$|Z|^2 = \varepsilon(Z)g^{jh}Z_j Z_h, \tag{1.37}$$

where $\varepsilon(Z) = \pm 1$ according as the quadratic form $g^{jh}Z_j Z_h$ is positive or negative. In particular, whenever $X^j$, $Z_j$ are related according to (1.33) or (1.35), one has $\varepsilon(X) = \varepsilon(Z)$ and $|X| = |Z|$ as a direct consequence of (1.36).

In conclusion we note that the constructions described above give rise to the following definition. With any two non-null contravariant vectors $X^j$, $Y^j$ in $T_n(P)$ we associate the following real function:

$$\cos(X, Y) = \frac{(X \cdot Y)}{|X \| Y|}, \tag{1.38}$$

in which the inner product is given by (1.20). This function is regarded as a cosine in the classical sense of the word since it refers to the "angle" defined by two vectors in the tangent space $T_n(P)$, which, as we have seen, is a Euclidean space whenever the metric is positive definite. In particular, the vectors $X$, $Y$ are said to be *normal* (or orthogonal) provided that

$$(X \cdot Y) = 0. \tag{1.39}$$

Because of the symmetry of the metric tensor $g_{jh}$, this condition of normality is a symmetric one. With the aid of the "unit hyperspheres" in $T_n(P)$ the cosine (1.38) may be interpreted geometrically in terms of adjacent sides of a triangle as in elementary Euclidean geometry.

## 7.2  GEODESICS

It was shown in Section 3.5 that one can always construct a unique connection on a differentiable manifold $X_n$ whenever a symmetric class $C^2$ tensor field $a_{hj}(x)$ is given, provided that the determinant $a = \det(a_{hj})$ is nonvanishing. The metric tensor $g_{hj}(x)$ of a Riemannian space $V_n$ satisfies the conditions under which this construction is possible, and consequently the metric defines a connection on $V_n$, which we shall now use consistently.

In accordance with (3.5.15) and (3.5.20) we can write down the Christoffel symbols of the first and second kinds; however, since we wish to emphasize that these symbols are defined with respect to the given metric we henceforth use the following notation:

$$\gamma_{hlk} = \frac{1}{2}\left(\frac{\partial g_{lk}}{\partial x^h} + \frac{\partial g_{hl}}{\partial x^k} - \frac{\partial g_{hk}}{\partial x^l}\right), \tag{2.1}$$

and

$$\gamma_{h\,k}^{\ j} = g^{jl}\gamma_{hlk} = \frac{1}{2}\,g^{jl}\left(\frac{\partial g_{lk}}{\partial x^h} + \frac{\partial g_{hl}}{\partial x^k} - \frac{\partial g_{hk}}{\partial x^l}\right), \tag{2.2}$$

these quantities representing the Christoffel symbols of the first and second kinds, respectively.

In the sequel, unless otherwise stated, the processes of covariant differentiation are defined with respect to the connection (2.2). Partial covariant differentiation with respect to the positional coordinate $x^k$ will again be denoted by a vertical stroke followed by the subscript $k$, and the operator $D$ will have the same meaning as before. In particular, it follows from (3.5.26) and (3.5.27) that the covariant derivative of the metric tensor vanishes identically:

$$g_{jh|k} = 0, \qquad g^{jh}{}_{|k} = 0, \tag{2.3}$$

these identities usually being referred to as *Ricci's lemma*. Also, for any displacement $dx^k$ in $V_n$ we have

$$Dg_{jh} = g_{jh|k}\,dx^k = 0 \tag{2.4}$$

identically.

A contravariant vector with components $X^j$ is once more said to be transported by *parallel displacement* from the point $P(x^k)$ to a neighboring point $Q(x^k + dx^k)$, if

$$DX^j = dX^j + \gamma_{h\,k}^{\ j}X^h\,dx^k = 0$$

for the given displacement $dx^k$. In particular, under these circumstances, we have

$$\begin{aligned}
d|X|^2 = D|X|^2 &= \varepsilon(X)D[g_{hj}X^hX^j] \\
&= \varepsilon(X)[(Dg_{hj})X^hX^j + g_{hj}(DX^h)X^j + g_{hj}X^h(DX^j)],
\end{aligned}$$

which vanishes by virtue of (2.4) and the condition that $DX^j = 0$. Thus $|X|^2 = $ constant: this implies that *the length of any vector remains unchanged under parallel displacement*.

As an example of the application of this process let us consider the line-element

$$ds^2 = g_{hj}\,dx^h\,dx^j \tag{2.5}$$

along a differentiable curve $x^j = x^j(s)$ referred to its arc length $s$ as parameter. Clearly the tangent vector $x'^j = dx^j/ds$ is a unit vector:

$$g_{hj}x'^hx'^j = 1. \tag{2.6}$$

Thus, according to the usual rules of covariant differentiation, we have

$$0 = \frac{d}{ds}(g_{hj}x'^hx'^j) = \frac{D}{Ds}(g_{hj}x'^hx'^j) = 2g_{hj}x'^h\frac{Dx'^j}{Ds}, \tag{2.7}$$

where, in the last step, we have used (2.4) and the symmetry of $g_{hj}$. This result indicates that the vector $Dx'^j/Ds$ is normal to the tangent vector $x'^j$. Let us denote by $\rho^{-1}$ the length of the vector $Dx'^j/Ds$:

$$\frac{1}{\rho} = \left|\frac{Dx'^j}{Ds}\right|. \tag{2.8}$$

Thus $\rho(Dx'^j/Ds)$ is a unit vector, to be denoted by $n^j$, which is normal to the curve $C$:

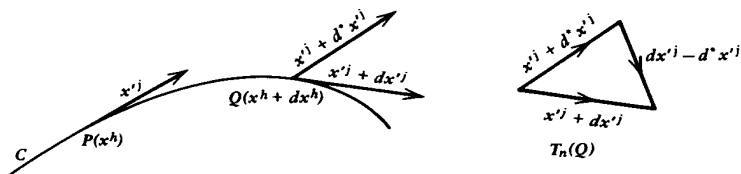$$\frac{Dx'^j}{Ds} = \frac{1}{\rho}\,n^j. \tag{2.9}$$

**Fig. 10**

We shall refer to $n^j$ as the *principal normal* to the curve $C$, and (2.9) is obviously a generalization of the first Frenet formula for a curve in a Euclidean space.

The following heuristic argument will serve to illustrate the meaning of the invariant defined by (2.8). Let us consider two neighboring points $P(x^h)$, $Q(x^h + dx^h)$ on the curve $C$. The unit tangent vectors at $P$ and $Q$ are respectively $x'^j$ and $x'^j + dx'^j$, where $dx'^j = (dx'^j/ds)\,ds = (d^2x^j/ds^2)\,ds$, which is obtained by repeated differentiation of the equations $x^j = x^j(s)$ defining $C$. Now let us transfer the unit vector $x'^j$ at $P$ by parallel displacement to $Q$ (see Figure 10). This yields the vector $x'^j + d^*x'^j$ at $Q$, where $d^*x'^j$ is given by

$$d^*x'^j + \gamma_h{}^j{}_k x'^h\, dx^k = 0 \qquad (2.10)$$

in accordance with the definition of parallel displacement of $x'^j$. Furthermore, since parallel displacement preserves lengths, the vector $x'^j + d^*x'^j$ at $Q$ is a unit vector. In terms of the Euclidean metric of the tangent space $T_n(Q)$ the angle $d\theta$ between the vectors $x'^j + d^*x'^j$, $x'^j + dx'^j$ is therefore defined to be

$$d\theta = |dx'^j - d^*x'^j|,$$

or, if we invoke (2.10),

$$d\theta = |dx'^j + \gamma_h{}^j{}_k x'^h\, dx^k| = |Dx'^j|.$$

Thus (2.8) is equivalent to

$$\frac{1}{\rho} = \frac{d\theta}{ds}, \qquad (2.11)$$

and accordingly we interpret $\rho^{-1}$ as the *curvature of the curve $C$* (see, e.g., Stoker [1]).

The arc length of any curve $C: x^j = x^j(t)$ between two points $P_1$ and $P_2$ corresponding to parameter values $t_1$ and $t_2$, respectively, can be calculated from the integral (1.2). We now discuss the conditions which the curve must satisfy in order that it afford an extreme value to this integral in the sense of Section 6.1. Since we must allow for the possibility that $F(x^j, \dot{x}^j)$ may vanish, we are obliged to follow a somewhat indirect approach.

For an arbitrary parameter $t$, we may write, using (1.15),

$$f(x^j, \dot{x}^j) = \tfrac{1}{2}F^2(x^j, \dot{x}^j) = \tfrac{1}{2}g_{hj}\dot{x}^h\dot{x}^j, \qquad (2.12)$$

so that, by (2.5),

$$f(x^j, \dot{x}^j) = \frac{1}{2}F^2(x^j, \dot{x}^j) = \frac{1}{2}\left(\frac{ds}{dt}\right)^2. \qquad (2.13)$$

Clearly

$$F\frac{\partial F}{\partial \dot{x}^j} = \frac{\partial f}{\partial \dot{x}^j}, \qquad F\frac{\partial F}{\partial x^j} = \frac{\partial f}{\partial x^j}, \qquad (2.14)$$

and

$$\frac{dF}{dt}\frac{\partial F}{\partial \dot{x}^j} + F\frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^j}\right) = \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}^j}\right). \qquad (2.15)$$

Thus, in terms of the definition (6.1.18), we have

$$FE_j(F) = E_j(f) - \frac{dF}{dt}\frac{\partial F}{\partial \dot{x}^j}. \qquad (2.16)$$

It also follows from (2.12) that

$$\frac{\partial f}{\partial \dot{x}^j} = g_{hj}\dot{x}^h, \qquad \frac{\partial f}{\partial x^j} = \frac{1}{2}\frac{\partial g_{hk}}{\partial x^j}\dot{x}^h\dot{x}^k,$$

so that

$$E_j(f) = \frac{d}{dt}\left(\frac{\partial f}{\partial \dot{x}^j}\right) - \frac{\partial f}{\partial x^j} = g_{hj}\ddot{x}^h + \frac{\partial g_{hj}}{\partial x^k}\dot{x}^h\dot{x}^k - \frac{1}{2}\frac{\partial g_{hk}}{\partial x^j}\dot{x}^h\dot{x}^k,$$

or, if we use (2.1) and (2.2),

$$E_j(f) = g_{hj}\ddot{x}^h + \gamma_{hjk}\dot{x}^h\dot{x}^k = g_{hj}(\ddot{x}^h + \gamma_l{}^h{}_k \dot{x}^l\dot{x}^k),$$

and hence

$$E_j(f) = g_{hj}\frac{D\dot{x}^h}{Dt}. \qquad (2.17)$$

The identity (2.16) can therefore be written in the form:

$$FE_j(F) = g_{hj}\frac{D\dot{x}^h}{Dt} - \frac{dF}{dt}\frac{\partial F}{\partial \dot{x}^j}. \qquad (2.18)$$

A curve $C: x^j = x^j(t)$ is said to be a *geodesic of the Riemannian space* $V_n$ if it is a solution of the Euler-Lagrange equation

$$E_j(F) = 0. \qquad (2.19)$$

If $F(x^j, \dot{x}^j) \neq 0$ along a geodesic, it follows from (2.18) that the latter satisfies the condition

$$g_{hj}\frac{D\dot{x}^h}{Dt} = \frac{dF}{dt}\frac{\partial F}{\partial \dot{x}^j}. \tag{2.20}$$

In particular, if the parameter $t$ is identified with the arc length $s$ of the curve, it is inferred from (2.13) that $F^2 = 1$, so that $dF/ds = 0$, in which case (2.20) reduces to

$$g_{hj}\frac{Dx'^h}{Ds} = 0,$$

or

$$\frac{Dx'^j}{Ds} = \frac{d^2x^j}{ds^2} + \gamma_h{}^j{}_k\frac{dx^h}{ds}\frac{dx^k}{ds} = 0. \tag{2.21}$$

This is the explicit form of the system of $n$ second-order ordinary differential equations to be satisfied by the geodesics of $V_n$ when referred to the arc length $s$ as parameter (it being assumed that $F \neq 0$). The equation (2.21) clearly shows that the geodesics of $V_n$ are *autoparallel curves*, that is, successive tangents are obtained from each other by parallel displacement along the curve. Moreover, it is evident from (2.8) that *geodesics are characterized by the vanishing of their curvature $\rho^{-1}$*. These phenomena are direct analogues of the corresponding properties of straight lines of Euclidean geometry; in fact, it follows from equation (2.5.25) that the differential equations characterizing straight lines in a Euclidean space referred to curvilinear coordinates have precisely the same formal structure as the equations (2.21) satisfied by the geodesics of $V_n$.

Instead of considering extreme values of the integral (1.2), one may consider the integral

$$I_f(C) = \int_C^{t_2}{}_{t_1} f(x^j, \dot{x}^j)\,dt. \tag{2.22}$$

A curve satisfying the Euler-Lagrange equations

$$E_j(f) = 0 \tag{2.23}$$

associated with (2.22) will be called an *f-extremal*. According to (2.17) an *f*-extremal is given by

$$\frac{D\dot{x}^h}{Dt} = 0. \tag{2.24}$$

This system of second-order ordinary differential equations always posseses a first integral. For, from (2.12) it follows that

$$2F\frac{dF}{dt} = \frac{d}{dt}(g_{hj}\dot{x}^h\dot{x}^j) = \frac{D}{Dt}(g_{hj}\dot{x}^h\dot{x}^j) = 2g_{jh}\dot{x}^j\frac{D\dot{x}^h}{Dt},$$

where, in the last step, we have used (2.4) and the symmetry of $g_{jh}$. Thus if (2.24) is satisfied, we have

$$F\frac{dF}{dt} = 0,$$

or

$$F(x^j, \dot{x}^j) = c, \tag{2.25}$$

where $c$ is some constant.

If $c \neq 0$, it follows from (2.13) that

$$\left(\frac{ds}{dt}\right)^2 = c^2, \tag{2.26}$$

so that $s = \pm ct + b$. Thus, for an $f$-extremal the parameter $t$—which was supposed to be arbitrary—turns out to be a linear function of the arc length $s$ whenever $F \neq 0$. However, under these circumstances it follows from (2.16) and (2.23) that $E_j(F) = 0$. *Accordingly an f-extremal satisfying the condition $F(x^j, \dot{x}^j) \neq 0$ is simply a geodesic of $V_n$.* Moreover, it should be observed that if $F \neq 0$ at an initial point $t = t_0$ of the geodesic, this condition will persist for all values of $t$ by virtue of (2.25).

On the other hand, if $c = 0$, the relation (2.26) does not determine the parameter $t$ in any way. Because of (2.24) and (2.25) an $f$-extremal of this kind is characterized by the relations

$$\frac{D\dot{x}^j}{Dt} = 0 \quad \text{and} \quad g_{hj}\dot{x}^h\dot{x}^j = 0. \tag{2.27}$$

A curve satisfying these conditions is called a *null geodesic of the Riemannian space $V_n$*. It should be remarked, however, that the formal structure of the first of these conditions is *not* independent of the choice of the parameter $t$. In fact, under a parameter transformation $\tau = \tau(t)$, with $d\tau/dt \neq 0$, the relations (2.27) become

$$\frac{d^2x^j}{d\tau^2} + \gamma_h{}^j{}_k\frac{dx^h}{d\tau}\frac{dx^k}{d\tau} = -\frac{dx^j}{d\tau}\left(\frac{d^2\tau}{dt^2}\right)\left(\frac{dt}{d\tau}\right)^2 \quad \text{and} \quad g_{hj}\frac{dx^h}{d\tau}\frac{dx^j}{d\tau} = 0. \tag{2.28}$$

Conversely, if a curve $C: x^j = x^j(\sigma)$, referred to some parameter $\sigma$, satisfies the conditions

$$\frac{d^2x^j}{d\sigma^2} + \gamma_h{}^j{}_k\frac{dx^h}{d\sigma}\frac{dx^k}{d\sigma} = g(\sigma)\frac{dx^j}{d\sigma} \quad \text{and} \quad g_{hj}\frac{dx^h}{d\sigma}\frac{dx^j}{d\sigma} = 0, \tag{2.29}$$

where $g(\sigma)$ is a given function of $\sigma$, it is easily inferred that $C$ is a null geodesic. For, if one defines a parameter $t$ by writing

$$t = \int \left[ \exp\left( \int g(\sigma) \, d\sigma \right) \right] d\sigma, \tag{2.30}$$

one has

$$\frac{dt}{d\sigma} = \exp\left( \int g(\sigma) \, d\sigma \right), \qquad \frac{d^2t}{d\sigma^2} = \frac{dt}{d\sigma} g(\sigma), \tag{2.31}$$

while (2.29) may be transformed to read

$$\left[ \frac{d^2x^j}{dt^2} + \gamma_h{}^j{}_k \frac{dx^h}{dt} \frac{dx^k}{dt} \right] \left( \frac{dt}{d\sigma} \right)^2 + \frac{dx^j}{dt} \frac{d^2t}{d\sigma^2} = g(\sigma) \frac{dx^j}{dt} \frac{dt}{d\sigma},$$

and, because of the second relation in (2.31), this reduces (2.29) directly to the form (2.27).

We have therefore established the following result: *if a curve $C: x^j = x^j(\sigma)$ of $V_n$ with $F(x^j, \dot{x}^j) = 0$ is such that the covariant derivative $D(dx^j/d\sigma)$ of its tangent vector $dx^j/d\sigma$ coincides in direction with this tangent vector, then $C$ is a null geodesic of $V_n$.* In this formulation the parameter $\sigma$ is arbitrary except for the condition $\sigma \neq s$. This characterization of null geodesics is therefore parameter-invariant and of a purely geometrical nature.

The argument leading to (2.25) can be applied to geodesics without reference to $f$-extremals; the equations (2.19) or (2.21) always possess the first integral (2.25). Thus, any geodesic issuing from a point $P(x^h_{(0)})$ of $V_n$, with initial direction $\dot{x}^j_{(0)}$ such that $F(x^j_{(0)}, \dot{x}^j_{(0)}) > 0$, has the property that $F(x^j, \dot{x}^j) > 0$ for all values of $t$. Similarly, if the initial direction of an $f$-extremal is a null direction, the curve will be a null geodesic.

In order to be able to distinguish between geodesics which afford maximum or minimum values to the length integral, we have to invoke the Weierstrass excess function. To this end we put $t = s$ in the integral (2.22), which is consistent with (2.26) provided that null geodesics are excluded from this discussion. From (2.13) we then have

$$I_f(C) = \frac{1}{2} \int_{s_1}^{s_2} ds, \tag{2.32}$$

so that, apart from the factor $\frac{1}{2}$, this integral does in fact represent the arc length of the curve $C$ between the points corresponding to the parameter values $s_1$ and $s_2$. Since

$$\frac{\partial^2 f}{\partial \dot{x}^h \partial \dot{x}^j} = g_{hj},$$

the expression (6.2.47) for the excess function associated with the integral (2.22) or (2.32) is simply

$$E(t, x^h, \psi^h, \dot{x}^h) = \tfrac{1}{2} g_{hj}(\dot{x}^h - \psi^h)(\dot{x}^j - \psi^j). \tag{2.33}$$

If the metric is positive definite this quadratic form is positive definite, so that $E \geq 0$. From the general theory of Section 6.2 it then follows that if a geodesic affords an extreme value to the length integral, this extremum is a *minimum*.

If the metric is indefinite, the situation is more complicated, and accordingly we restrict our consideration to the case $n = 2$ with the metric (1.31). Let us consider a geodesic whose unit tangent vectors are given by $\psi^j(s)$, such that $\varepsilon(\psi) = +1$ (a property which, as we have seen, persists along any geodesic if it is satisfied at a single point). As admissible curves of comparison we admit solely such curves $x^j = x^j(s)$ whose unit tangent vectors $x'^j(s)$ also satisfy the condition $\varepsilon(x') = +1$. Again (2.32) represents the length integral (apart from the factor $+\frac{1}{2}$). However, if $\varepsilon(\psi) = +1$, $\varepsilon(x') = +1$ for two unit vectors in any tangent space $T_2(P)$, we have $\varepsilon(x' - \psi) = -1$ (see Figure 9 of Section 7.1). Under these circumstances it follows from (2.33) that $E(t, x^h, \psi^h, \dot{x}^h) \leq 0$. Thus, subject to the above restrictions, we infer that if our geodesic affords an extreme value to the length integral, this extremum is a *maximum*.

These conclusions are in agreement with the observations of Section 7.1 concerning the triangle inequality and the reversed triangle inequality.

## 7.3 CURVATURE THEORY OF RIEMANNIAN SPACES

In Section 3.6 it was shown how one can associate a curvature tensor $K_l{}^j{}_{hk}$ with a given affine connection field $\Gamma_h{}^j{}_k$ on a differentiable manifold $X_n$. In the case of a Riemannian space, for which the connection may be identified with the Christoffel symbols $\gamma_h{}^j{}_k$, one would naturally use the curvature tensor defined in terms of this special connection. This tensor will henceforth be denoted by $R_l{}^j{}_{hk}$, and, according to the formula (3.6.8), it is given by[†]

$$R_l{}^j{}_{hk} = \frac{\partial \gamma_l{}^j{}_h}{\partial x^k} - \frac{\partial \gamma_l{}^j{}_k}{\partial x^h} + \gamma_m{}^j{}_k \gamma_l{}^m{}_h - \gamma_m{}^j{}_h \gamma_l{}^m{}_k. \tag{3.1}$$

The theory developed in Sections 3.6–3.7 may clearly be applied directly to the theory of Riemannian manifolds. In particular, since the Christoffel symbols $\gamma_h{}^j{}_k$ are symmetric in their subscripts $h$ and $k$ [as is immediately

[†] In the case of $n = 4$, there exist various computer programs which algebraically compute the curvature tensor corresponding to a particular metric tensor; see Barton and Fitch [1].

evident from (2.2)], the corresponding torsion tensor vanishes identically, and consequently the resulting formalism is simplified considerably. In particular, the general formula (3.6.11) concerning an interchange of the orders of repeated partial covariant differentiation will now reduce to

$$T^{j_1 \cdots j_r}{}_{l_1 \cdots l_s | h | k} - T^{j_1 \cdots j_r}{}_{l_1 \cdots l_s | k | h} = \sum_{\alpha=1}^{r} R_m{}^{j_\alpha}{}_{hk} T^{j_1 \cdots j_{\alpha-1} m j_{\alpha+1} \cdots j_r}{}_{l_1 \cdots l_s}$$

$$- \sum_{\beta=1}^{s} R_{l_\beta}{}^{m}{}_{hk} T^{j_1 \cdots j_r}{}_{l_1 \cdots l_{\beta-1} m l_{\beta+1} \cdots l_s}. \tag{3.2}$$

Special cases of this formula are the following:

$$X^j_{|h|k} - X^j_{|k|h} = X^l R_l{}^j{}_{hk} \tag{3.3}$$

for any class $C^2$ contravariant vector field, while for any class $C^2$ type $(0, 2)$ tensor field

$$A_{jl|h|k} - A_{jl|k|h} = -A_{ml} R_j{}^m{}_{hk} - A_{jm} R_l{}^m{}_{hk}. \tag{3.4}$$

The tensor $R_l{}^j{}_{hk}$ is of type $(1, 3)$. With the aid of our metric tensor $g_{jm}$ we can construct a type $(0, 4)$ tensor by writing

$$R_{lmhk} = g_{jm} R_l{}^j{}_{hk}, \tag{3.5}$$

which is usually referred to as the covariant curvature tensor. The latter satisfies various identities which we shall now enumerate in a systematic manner. From (3.8.1) and (3.8.7) we know that

$$R_l{}^j{}_{hk} = -R_l{}^j{}_{kh}, \tag{3.6}$$

and

$$R_l{}^j{}_{hk} + R_h{}^j{}_{kl} + R_k{}^j{}_{lh} = 0. \tag{3.7}$$

When these relations are multiplied by $g_{jm}$, it is found in terms of the notation (3.5) that

$$R_{lmhk} = -R_{lmkh}, \tag{3.8}$$

and

$$R_{lmhk} + R_{hmkl} + R_{kmlh} = 0. \tag{3.9}$$

A further identity is obtained when we identify the tensor $A_{jl}$ in (3.4) with the metric tensor $g_{jl}$. Because of Ricci's lemma (2.3) one has

$$0 = -g_{ml} R_j{}^m{}_{hk} - g_{jm} R_l{}^m{}_{hk},$$

or, using (3.5) once more,

$$R_{jlhk} = -R_{ljhk}. \tag{3.10}$$

Thus the covariant curvature tensor is skew-symmetric in the first two subscripts as well as in the last two subscripts.

The identities (3.8), (3.9), and (3.10) also imply the useful relation

$$R_{ljhk} = R_{hklj}. \tag{3.11}$$

This is easily verified as follows. From (3.9) and (3.10) we have

$$R_{ljhk} = -R_{lhkj} - R_{lkjh}, \tag{3.12}$$

while

$$R_{lhkj} = -R_{hlkj} = R_{hkjl} + R_{hjlk}, \qquad R_{lkjh} = -R_{kljh} = R_{kjhl} + R_{khlj}.$$

The last two expressions are substituted in (3.12), the skew-symmetry relations (3.8) and (3.10) being taken into account once more. This gives

$$R_{ljhk} = 2R_{hklj} - R_{hjlk} - R_{kjhl} = 2R_{hklj} + R_{ljkh},$$

where, in the last step, we have used (3.9) once more. This establishes (3.11).

For many applications of the theory of curvature it is necessary to ascertain the number $N$ of independent components of the tensor (3.5). In order to determine $N$ let us suspend the summation convention for the moment, and let $l$, $j$, $h$, $k$ denote distinct integers. There are five cases to be considered: (1) all four indices identical; (2) three indices identical; (3) two pairs of identical indices; (4) two indices identical, two distinct; (5) all four indices distinct. Cases 1 and 2 do not yield any independent components since the latter are identically zero under these circumstances. For case 3 we note that $R_{ljlj}$ has $\binom{n}{2}$ independent components, while any components with different positions of $l$ and $j$, for example, $R_{jllj}$, are expressible in terms of $R_{ljlj}$ by virtue of (3.10). With regard to case 4 we observe that $R_{ljlk}$ has $n\binom{n-1}{2}$ independent components, while all other components of this kind, for example, $R_{lklj}$ is expressible in terms of $R_{ljlk}$ because of (3.11). Finally, for case 5 we note that $R_{ljhk}$ has $\binom{n}{4}$ independent components; however, $R_{lhkj}$ is not expressible in terms of $R_{ljhk}$, and accordingly this case yields $2\binom{n}{4}$ independent components, all others being expressible in terms of $R_{ljhk}$ and $R_{lhkj}$. Hence

$$N = \binom{n}{2} + n\binom{n-1}{2} + 2\binom{n}{4} = \frac{1}{12} n^2(n^2 - 1).$$

For instance, when $n = 2$, there is only one independent component, say, $R_{1212}$, all other components being zero or $\pm R_{1212}$. Incidentally, it can be shown that the Gaussian curvature of a $V_2$ is proportional to $R_{1212}$.

It is sometimes useful to have at one's disposal an explicit representation of the curvature tensor (3.5) in terms of derivatives of the metric tensor.

On substituting (3.1) in (3.5) we find that

$$R_{jlhk} = g_{lm} R_j{}^m{}_{hk} = g_{lm} \frac{\partial \gamma_j{}^m{}_h}{\partial x^k} - g_{lm} \frac{\partial \gamma_j{}^m{}_k}{\partial x^h} + g_{lm} \gamma_p{}^m{}_k \gamma_j{}^p{}_h - g_{lm} \gamma_p{}^m{}_h \gamma_j{}^p{}_k$$

$$= \frac{\partial \gamma_{jlh}}{\partial x^k} - \frac{\partial \gamma_{jlk}}{\partial x^h} - \frac{\partial g_{lm}}{\partial x^k} \gamma_j{}^m{}_h + \frac{\partial g_{lm}}{\partial x^h} \gamma_j{}^m{}_k + \gamma_{mlk} \gamma_j{}^m{}_h - \gamma_{mlh} \gamma_j{}^m{}_k.$$

(3.13)

However, from (2.1) we have

$$\frac{\partial \gamma_{jlh}}{\partial x^k} = \frac{1}{2} \frac{\partial}{\partial x^k} \left( \frac{\partial g_{lh}}{\partial x^j} + \frac{\partial g_{lj}}{\partial x^h} - \frac{\partial g_{jh}}{\partial x^l} \right) = \frac{1}{2} \left( \frac{\partial^2 g_{lh}}{\partial x^k \partial x^j} + \frac{\partial^2 g_{lj}}{\partial x^k \partial x^h} - \frac{\partial^2 g_{jh}}{\partial x^k \partial x^l} \right),$$

so that, since $g_{lj}(x)$ is of class $C^2$,

$$\frac{\partial \gamma_{jlh}}{\partial x^k} - \frac{\partial \gamma_{jlk}}{\partial x^h} = \frac{1}{2} \left( \frac{\partial^2 g_{lh}}{\partial x^k \partial x^j} - \frac{\partial^2 g_{lk}}{\partial x^h \partial x^j} + \frac{\partial^2 g_{jk}}{\partial x^h \partial x^l} - \frac{\partial^2 g_{jh}}{\partial x^k \partial x^l} \right).$$

(3.14)

Also, because of Ricci's lemma, we have

$$\frac{\partial g_{lm}}{\partial x^k} = g_{pm} \gamma_l{}^p{}_k + g_{lp} \gamma_m{}^p{}_k = \gamma_{lmk} + \gamma_{mlk},$$

(3.15)

with a similar formula for $\partial g_{lm}/\partial x^h$. When these expressions, together with (3.14), are substituted in (3.13), several obvious simplifications being effected, we obtain the following formula for the covariant curvature tensor:

$$R_{jlhk} = \frac{1}{2} \left( \frac{\partial^2 g_{lh}}{\partial x^k \partial x^j} - \frac{\partial^2 g_{lk}}{\partial x^h \partial x^j} + \frac{\partial^2 g_{jk}}{\partial x^h \partial x^l} - \frac{\partial^2 g_{jh}}{\partial x^k \partial x^l} \right) + \gamma_{lmh} \gamma_j{}^m{}_k - \gamma_{lmk} \gamma_j{}^m{}_h.$$

(3.16)

It is evident that the curvature tensor is linear in the second derivatives of the metric tensor; it is nonlinear in the first derivatives. Furthermore, the identities (3.8), (3.9), (3.10), and (3.11) may be deduced directly from (3.16) by inspection.

These identities are of a purely algebraic character. From the theory of Section 3.8 it should be recalled, however, that the curvature tensor of an affinely connected manifold satisfies a set of *differential* identities, namely, the Bianchi identities. Clearly this must be true in particular for the curvature tensor (3.1) of our Riemannian space; moreover, since the Christoffel symbols define a symmetric connection, we may infer directly from (3.8.12) that

$$R_j{}^p{}_{hk|m} + R_j{}^p{}_{mh|k} + R_j{}^p{}_{km|h} = 0.$$

(3.17)

From (3.5) and Ricci's lemma we also have

$$R_{jlhk|m} = g_{lp} R_j{}^p{}_{hk|m},$$

and hence multiplication of (3.17) by $g_{lp}$ yields

$$R_{jlhk|m} + R_{jlmh|k} + R_{jlkm|h} = 0,$$

(3.18)

which are the Bianchi identities for the covariant curvature tensor. In passing we note that the symmetry of the Christoffel symbols permits the introduction of normal coordinates, by means of which (3.17) can be deduced almost directly from (3.1).

By analogy with (3.8.13) we may also define the *Ricci tensor* $R_{lh}$ by contraction of the curvature tensor (3.1), namely, by writing

$$R_{lh} = R_l{}^j{}_{hj} = g^{pj} R_{lphj},$$

(3.19)

so that

$$R_{lh} = \frac{\partial \gamma_l{}^j{}_h}{\partial x^j} - \frac{\partial \gamma_l{}^j{}_j}{\partial x^h} + \gamma_m{}^j{}_j \gamma_l{}^m{}_h - \gamma_m{}^j{}_h \gamma_l{}^m{}_j.$$

(3.20)

Now, applying the formula (4.1.31) to the metric tensor $g_{hj}$ we have the identity

$$\gamma_l{}^j{}_j = \frac{\partial}{\partial x^l} (\ln \sqrt{g}),$$

(3.21)

where

$$g = |\det(g_{hj})|.$$

(3.22)

Thus (3.20) can be written in the form

$$R_{lh} = \frac{\partial \gamma_l{}^j{}_h}{\partial x^j} - \frac{\partial^2}{\partial x^h \partial x^l} (\ln \sqrt{g}) + \frac{\partial}{\partial x^m} (\ln \sqrt{g}) \gamma_l{}^m{}_h - \gamma_m{}^j{}_h \gamma_l{}^m{}_j,$$

(3.23)

from which it is clear that the Ricci tensor is symmetric when constructed by means of the Christoffel symbols:

$$R_{lh} = R_{hl}.$$

(3.24)

This conclusion is in agreement with the general theory discussed at the end of Section 3.8.

By analogy with (3.5) we shall use the notation

$$R_h^j = g^{lj} R_{lh},$$

(3.25)

which suggests the further contraction

$$R = R_j^j = g^{lj} R_{lj}.$$

(3.26)

Clearly the quantity thus defined is an invariant, which will be referred to as the *curvature scalar*. By means of the latter and the Ricci tensor one may

construct the so-called *Einstein tensor*:

$$G_h^j = R_h^j - \tfrac{1}{2}\delta_h^j R. \tag{3.27}$$

This tensor is of fundamental importance, which is due to the fact that *its divergence vanishes identically*:

$$G_{h|j}^j = 0. \tag{3.28}$$

The proof of this assertion is based on the Bianchi identity (3.18). Let us multiply the latter by $g^{lk}$, noting (3.19) and Ricci's lemma, while (3.8) is applied to the third term on the left-hand side. This yields

$$R_{jh|m} + g^{lk}R_{jlmh|k} - R_{jm|h} = 0.$$

This relation is now multiplied by $g^{jh}$, the definition (3.26) being taken into account, while (3.10) and (3.19) are used to simplify the term in the middle. It is thus found that

$$R_{|m} - g^{lk}R_{lm|k} - g^{jh}R_{jm|h} = 0,$$

and by means of (3.25) and (3.27) this can immediately be reduced to the assertion (3.28).

By virtue of (2.3) the identity (3.28) can also be expressed in the form

$$G^{jh}{}_{|j} = 0, \tag{3.29}$$

where

$$G^{jh} = R^{jh} - \tfrac{1}{2}g^{jh}R \tag{3.30}$$

and

$$R^{jh} = g^{hk}R_k^j. \tag{3.31}$$

Some of the most remarkable theorems concerning the curvature theory of Riemannian manifolds refer to special Riemannian spaces, a few of which are now considered. Of special importance in the general theory of relativity are the so-called Einstein spaces, which are characterized by the property that the Ricci tensor $R_{hl}$ is proportional to the metric tensor $g_{hl}$, that is,

$$R_{hl} = \lambda g_{hl}, \tag{3.32}$$

which is possible by virtue of the symmetry relation (3.24). The factor $\lambda$ of proportionality is obtained by multiplication of (3.32) by $g^{hl}$, the definition (3.26) being taken into account, which gives

$$R_{hl} = \frac{R}{n} g_{hl}, \tag{3.33}$$

or equivalently,

$$R_h^j = \frac{R}{n}\,\delta_h^j, \tag{3.34}$$

these being the defining equations of an *Einstein space*.†

The Einstein tensor (3.27) assumes a particularly simple form under these circumstances, namely,

$$G_h^j = \left(\frac{1}{n} - \frac{1}{2}\right)\delta_h^j R,$$

so that (3.28) yields

$$\left(\frac{1}{n} - \frac{1}{2}\right)R_{|h} = 0. \tag{3.35}$$

Therefore, *unless $n = 2$*,

$$R_{|h} = 0 \tag{3.36}$$

*for an Einstein space*, which implies that the curvature scalar is a constant for such spaces. In particular, therefore, the factor of proportionality in (3.33) is a constant. In view of (3.23) it is evident that (3.33) represents a system of $\frac{1}{2}n(n + 1)$ partial differential equations of the second order to be satisfied by the components $g_{hj}$ of the metric tensor, which are linear in the second derivatives of the $g_{hj}$, but nonlinear in the first derivatives.

The conclusion (3.36) is not valid when $n = 2$. This phenomenon is related to the fact that *every two-dimensional Riemannian space is an Einstein space*, an assertion which is easily verified as follows by means of the defining equation (3.19) for the Ricci tensor. Let us put $l = 1$, $h = 1$ in (3.19), after which we sum over $p$ and $j$ from 1 to 2, noting at the same time that the terms involving $R_{111j}$ and $R_{1p11}$ vanish identically as a result of (3.10) and (3.8), respectively. It is thus found that

$$R_{11} = g^{22}R_{1212}, \tag{3.37}$$

while similarly

$$R_{12} = R_{21} = -g^{21}R_{1212}, \qquad R_{22} = g^{11}R_{1212}. \tag{3.38}$$

However, for $n = 2$,

$$g^{hj} = \begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} = \begin{pmatrix} g_{22} & -g_{12} \\ -g_{21} & g_{11} \end{pmatrix}[\det(g_{lk})]^{-1},$$

† For an extensive study of four-dimensional Einstein spaces see Petrov [1].

so that the relations (3.37) and (3.38) are equivalent to

$$R_{hj} = g_{hj} \frac{R_{1212}}{\det(g_{lk})}. \tag{3.39}$$

This is of the form (3.33) with $n = 2$, where it is to be noted in passing that (3.26), (3.37), and (3.38) give

$$R = g^{11}R_{11} + 2g^{12}R_{12} + g^{22}R_{22} = (g^{11}g^{22} - 2(g^{12})^2 + g^{22}g^{11})R_{1212}$$
$$= 2 \det(g^{hj})R_{1212} = 2[\det(g_{hj})]^{-1}R_{1212}, \tag{3.40}$$

as required by consistency with (3.33).

A class of Riemannian manifolds which is even more special than that exemplified by Einstein spaces is represented by the so-called spaces of constant curvature. In order to define the latter it is necessary that we should first introduce the concept of sectional curvature of a Riemannian space. To this end let us consider a fixed point $P(x^h)$ of our $V_n$, and let us denote by $X^j$, $Y^j$ the components of two arbitrary linearly independent vectors in $T_n(P)$. These vectors determine a two-dimensional plane $T_2(P)$ of $T_n(P)$. The scalar defined by

$$K(x, X, Y) = \frac{R_{ljhk}X^lX^hY^jY^k}{(g_{lh}g_{jk} - g_{lk}g_{jh})X^lX^hY^jY^k}, \tag{3.41}$$

is called the *sectional* (or *Riemannian*) *curvature* of our $V_n$ at $P$ with respect to the 2-plane $T_2(P)$ (Riemann [1]). [This concept can be motivated geometrically as follows. The set of all geodesics of $V_n$ which pass through $P$ and are tangent to $T_2(P)$ at $P$ generates a uniquely defined two-dimensional subspace $V_2$ of $V_n$, which is said to be *geodesic* at $P$. It is possible to apply the techniques of the classical differential geometry of two-dimensional surfaces to this $V_2$, and accordingly its Gaussian curvature at $P$ can be evaluated. It may be shown (although this is not done here) that this Gaussian curvature is identical with the expression (3.41).]

If $U^j$, $V^j$ denote the components of two linearly independent vectors of $T_n(P)$ which are contained in $T_2(P)$, it is possible to write

$$U^j = a_1X^j + a_2Y^j, \qquad V^j = b_1X^j + b_2Y^j,$$

with $A = a_1b_2 - a_2b_1 \neq 0$. Because of the skew-symmetry properties (3.8) and (3.10) it is easily verified by direct calculation that, under these circumstances,

$$R_{ljhk}U^lU^hV^jV^k = A^2R_{ljhk}X^lX^hY^jY^k,$$

together with

$$(g_{lh}g_{jk} - g_{lk}g_{jh})U^lU^hV^jV^k = A^2(g_{lh}g_{jk} - g_{lk}g_{jh})X^lX^hY^jY^k.$$

From (3.41) it therefore follows that

$$K(x, X, Y) = K(x, U, V)$$

whenever the pairs $(X^j, Y^j)$, $(U^j, V^j)$ span the same 2-plane $T_2(P)$, from which it is evident that the sectional curvature depends solely on this 2-plane, and not on the choice of the vectors which span $T_2(P)$.

However, it may happen that at some point $P$ of $V_n$ the sectional curvature is the same for *all* 2-planes in $T_n(P)$. A point possessing this property is called *isotropic*. At an isotropic point the expression (3.41) assumes the same values for *all* pairs $(X^j, Y^j)$ of $T_n(P)$, in which case we denote $K(x, X, Y)$ by $k(x)$; moreover, it is obvious from (3.41) that a necessary and sufficient condition for $P$ to be isotropic is that

$$S_{ljhk}X^lX^hY^jY^k = 0 \tag{3.42}$$

for *any* pair of vectors $X^j$, $Y^j$, where, for the sake of brevity, we have put

$$S_{ljhk} = R_{ljhk} - k(x)(g_{lh}g_{jk} - g_{lk}g_{jh}). \tag{3.43}$$

By virtue of the quotient theorem the symmetric part of the coefficient of $X^lX^h$ in (3.42) must vanish:

$$(S_{ljhk} + S_{hjlk})Y^jY^k = 0,$$

and by the same token, the symmetric part of the coefficient of $Y^jY^k$ must be zero, that is,

$$S_{ljhk} + S_{lkhj} + S_{hjlk} + S_{hklj} = 0. \tag{3.44}$$

Now it is easily verified that the tensor $S_{ljhk}$ possesses precisely the same symmetry and skew-symmetry properties as the curvature tensor $R_{ljhk}$. It therefore follows from the counterpart of (3.11) that (3.44) reduces to

$$S_{ljhk} + S_{lkhj} = 0. \tag{3.45}$$

But from (3.9) we have

$$S_{lkhj} = -S_{lhjk} - S_{ljkh},$$

so that, with the aid of (3.8), one can write (3.45) as

$$2S_{ljhk} - S_{lhjk} = 0. \tag{3.46}$$

Also, since $S_{ljhk}$ is skew-symmetric in $h$ and $k$, it is evident from (3.45) that $S_{lkhj}$ is skew-symmetric in $h$ and $k$, so that (3.46) implies that

$$2S_{ljhk} = S_{lhjk} = -S_{ljhk},$$

which is possible if and only if $S_{ljhk} = 0$. By means of (3.43) it is inferred that *an isotropic point P of $V_n$ is characterized by the condition that the relations*

$$R_{ljhk} = k(x)(g_{lh}g_{jk} - g_{lk}g_{jh}) \qquad (3.47)$$

*be valid at P.*

If the Riemannian space $V_n$ consists entirely of isotropic points, that is, if the curvature tensor is everywhere expressible in the form (3.47), the manifold $V_n$ is said to be a *space of constant curvature*. The Ricci tensor of such a space is obtained directly from (3.47) by multiplication with $g^{jk}$, which gives

$$R_{lh} = (n - 1)k(x)g_{lh}. \qquad (3.48)$$

It therefore follows that a *space of constant curvature is an Einstein space*, and a comparison with (3.33) shows that the sectional curvature $k(x)$ is related to the curvature scalar $R$ according to

$$R = n(n - 1)k(x). \qquad (3.49)$$

Moreover, equation (3.36) indicates that for any Einstein space whose dimension exceeds 2, the scalar $R$ is a constant. It therefore follows from (3.49) that *for any Riemannian space $V_n$ of constant curvature, with $n > 2$, the sectional curvature $k(x)$ is a constant*. This result is generally known as *Schur's theorem*.

It should be remarked that the converse to the statement following (3.48) is not true, that is, an Einstein space need not be a space of constant curvature. Nor is an arbitrary two-dimensional Riemannian manifold necessarily a space of constant curvature. However, if a $V_2$ satisfies the condition (3.47) with $k(x) = $ constant, it exemplifies the so-called non-Euclidean geometries; the corresponding geometry is said to be elliptic or hyperbolic according as $k > 0$ or $k < 0$.

In conclusion, let us briefly glance at Riemannian manifolds whose curvature tensor vanishes identically. From the theorem at the end of Section 5.6 we infer, as a special case, that under these circumstances there exist coordinate systems in which the Christoffel symbols are identically zero. By virtue of the identity (3.15) it follows that the components of the metric tensor are constants in such coordinate systems, which is obviously characteristic of flat spaces. This conclusion may be summarized in the following theorem.

**THEOREM**

*In order that a Riemannian space be flat, it is necessary and sufficient that the components of its curvature tensor vanish identically.*

## 7.4  SUBSPACES OF A RIEMANNIAN MANIFOLD

As a special case of the theory developed in Section 5.7 we now consider a subspace $V_m$ of our $n$-dimensional Riemannian manifold $V_n$, (Eisenhart [1], Schouten [1], Rund [4]), this subspace being represented parametrically by the equations

$$x^j = x^j(u^\alpha) \qquad (j = 1, \ldots, n; \alpha = 1, \ldots, m; m < n). \qquad (4.1)$$

(Throughout this section Latin and Greek indices range from 1 to $n$ and from 1 to $m$, respectively, the summation convention being operative in respect of both sets of indices.) It is assumed that the functions $x^j(u^\alpha)$ are of class $C^3$, and that the rank of the matrix of the derivatives

$$B_\alpha^j = \frac{\partial x^j}{\partial u^\alpha} \qquad (4.2)$$

is maximal, namely, $m$. It will be recalled from Section 5.7 that, at each point $P$ of $V_m$, the $B_\alpha^j$ represent the components of $m$ linearly independent vectors in the tangent space $T_n(P)$ which span the tangent space $T_m(P)$ of $V_m$ at $P$.

A displacement in $V_n$ whose components $dx^j$ are expressible in the form

$$dx^j = B_\alpha^j \, du^\alpha \qquad (4.3)$$

is said to be tangential to $V_m$; the element of arc length $ds$ associated with such a displacement is given by

$$ds^2 = g_{hj} \, dx^h \, dx^j = g_{hj} B_\alpha^h B_\beta^j \, du^\alpha \, du^\beta. \qquad (4.4)$$

This clearly suggests the definition

$$g_{\alpha\beta}(u^\varepsilon) = g_{hj}(x^l(u^\varepsilon))B_\alpha^h B_\beta^j, \qquad (4.5)$$

so that, on $V_m$,

$$ds^2 = g_{\alpha\beta} \, du^\alpha \, du^\beta. \qquad (4.6)$$

[Again it is remarked that, strictly speaking, the quantities defined by (4.5) should be denoted by a symbol other than $g$; however, since in this context the subscripts are always either all Latin, or all Greek, no confusion is likely to arise.] Clearly the components (4.5) are class $C^3$ functions of their $m$ arguments $u^\varepsilon$, their symmetry being guaranteed by the symmetry of the metric tensor $g_{hj}$ of $V_n$. Moreover, in view of the fact that the $B_\alpha^h$ are type (0, 1) tensors relative to parameter transformations on $V_m$, the $g_{\alpha\beta}$ are components of a type (0, 2) tensor relative to such transformations; relative to coordinate transformations on $V_n$ they are simply scalars. Since the metric of $V_n$ is not assumed to be positive definite, it is possible that on certain subspaces $V_m$ there exist points for which $\det(g_{\alpha\beta}) = 0$. In order to exclude this

possibility and the complications resulting therefrom, we restrict ourselves to subspaces for which the condition

$$\det(g_{\alpha\beta}) \neq 0 \qquad (4.7)$$

is everywhere satisfied. Under these conditions we may regard the quantities (4.5) as the components of the *metric tensor* on $V_m$; the resulting metric of $V_m$ is said to be *induced* on $V_m$ by the metric of $V_n$. The subspace $V_m$ is a Riemannian manifold in its own right. The quantities $g_{\alpha\beta}$ are also called the coefficients of the *first fundamental form*.

The condition (4.7) guarantees the existence of a tensor $g^{\alpha\beta}$ inverse to the metric tensor:

$$g^{\alpha\beta} g_{\beta\varepsilon} = \delta^{\alpha}_{\varepsilon}, \qquad (4.8)$$

by means of which we construct the quantities

$$B^{\alpha}_{j} = g^{\alpha\varepsilon} g_{hj} B^{h}_{\varepsilon}. \qquad (4.9)$$

Because of (4.5) and (4.8) we have

$$B^{\alpha}_{j} B^{j}_{\beta} = g^{\alpha\varepsilon} g_{hj} B^{h}_{\varepsilon} B^{j}_{\beta} = g^{\alpha\varepsilon} g_{\varepsilon\beta} = \delta^{\alpha}_{\beta}; \qquad (4.10)$$

however, it should be noted that $B^{\alpha}_{j} B^{h}_{\alpha} \neq \delta^{h}_{j}$. Any element $Z^{j}$ of $T_{n}(P)$ is said to be *normal* to $V_m$ at $P$ if it satisfies the $m$ relations

$$g_{hj} Z^{h} B^{j}_{\alpha} = 0, \qquad (4.11)$$

since, under these conditions, the vector $Z^{j}$ is normal to each of the $m$ vectors $B^{j}_{\alpha}$ which span $T_{m}(P)$ in $T_{n}(P)$. Clearly one can always construct sets of $n - m$ linearly independent vectors in $T_{n}(P)$ which are normal to $V_m$ at $P$; however, unless $m = n - 1$, such sets are by no means uniquely defined. Moreover, it follows from (4.9) and (4.11) that

$$B^{\alpha}_{j} Z^{j} = 0 \qquad (4.12)$$

for any vector $Z^{j}$ normal to $V_m$.

The Christoffel symbols of $V_m$ may be defined precisely as in $V_n$ in terms of the derivatives of $g_{\alpha\beta}$, namely, as

$$\gamma_{\alpha}{}^{\lambda}{}_{\beta} = g^{\lambda\varepsilon} \gamma_{\alpha\varepsilon\beta} = \frac{1}{2} g^{\lambda\varepsilon} \left( \frac{\partial g_{\beta\varepsilon}}{\partial u^{\alpha}} + \frac{\partial g_{\varepsilon\alpha}}{\partial u^{\beta}} - \frac{\partial g_{\alpha\beta}}{\partial u^{\varepsilon}} \right). \qquad (4.13)$$

These three-index symbols determine a connection on $V_m$ in exactly the same way as the Christoffel symbols $\gamma_{h}{}^{j}{}_{k}$ define a connection on the embedding space $V_n$. By means of these two connections a process of mixed covariant differentiation as described in Section 5.7 is uniquely defined, this process being denoted once more by a pair of vertical bars followed by the appropriate subscript. Again, for quantities which are tensorial relative to param-

eter transformations on $V_m$, while being scalars relative to coordinate transformations in $V_n$, the mixed covariant derivative is identical with the usual covariant derivative on $V_m$. In particular, Ricci's lemma for $V_m$ is simply

$$g_{\alpha\beta\|\gamma} = 0. \qquad (4.14)$$

Also, since $g_{hj}(x)$ is defined on all of $V_n$ (and not merely on $V_m$), we may write

$$g_{hj\|\beta} = \frac{\partial g_{hj}}{\partial u^{\beta}} - g_{lj} \gamma_{h}{}^{l}{}_{k} B^{k}_{\beta} - g_{hl} \gamma_{j}{}^{l}{}_{k} B^{k}_{\beta} = \left( \frac{\partial g_{hj}}{\partial x^{k}} - g_{lj} \gamma_{h}{}^{l}{}_{k} - g_{hl} \gamma_{j}{}^{l}{}_{k} \right) B^{k}_{\beta} = g_{hj\|k} B^{k}_{\beta},$$

or, by Ricci's lemma for $V_n$,

$$g_{hj\|\beta} = 0. \qquad (4.15)$$

As in (5.7.32) let us now write

$$H_{\alpha}{}^{j}{}_{\beta} = B^{j}_{\alpha\|\beta} = (B_{\alpha}{}^{j}{}_{\beta} - \gamma_{\alpha}{}^{\lambda}{}_{\beta} B^{j}_{\lambda} + \gamma_{h}{}^{j}{}_{k} B^{h}_{\alpha} B^{k}_{\beta}), \qquad (4.16)$$

where we have used the notation

$$B_{\alpha}{}^{j}{}_{\beta} = \frac{\partial B^{j}_{\alpha}}{\partial u^{\beta}} = \frac{\partial^{2} x^{j}}{\partial u^{\alpha} \partial u^{\beta}} = B_{\beta}{}^{j}{}_{\alpha}. \qquad (4.17)$$

Because of the symmetry of both of the Christoffel symbols which appear in (4.16), it is evident that

$$H_{\alpha}{}^{j}{}_{\beta} = H_{\beta}{}^{j}{}_{\alpha}, \qquad (4.18)$$

it being recalled that the quantities (4.16) represent the components of $\frac{1}{2} m(m + 1)$ type (1, 0) tensors relative to coordinate transformations on $V_m$, and of $n$ type (0, 2) tensors relative to parameter transformations on $V_n$. Now the mixed covariant differentiation of (4.9) with respect to $u^{\beta}$ yields, in view of (4.14), (4.15), and (4.16):

$$B^{\alpha}_{j\|\beta} = g^{\alpha\lambda} g_{hj} H_{\lambda}{}^{h}{}_{\beta}. \qquad (4.19)$$

From (4.10) we obtain similarly

$$B^{\alpha}_{j\|\beta} B^{j}_{\varepsilon} + B^{\alpha}_{j} H_{\varepsilon}{}^{j}{}_{\beta} = 0,$$

and hence, by (4.18),

$$B^{\alpha}_{j\|\beta} B^{j}_{\varepsilon} = B^{\alpha}_{j\|\varepsilon} B^{j}_{\beta},$$

so that (4.19), taken together with (4.8), gives

$$g_{hj} H_{\alpha}{}^{h}{}_{\beta} B^{j}_{\varepsilon} = g_{hj} H_{\alpha}{}^{h}{}_{\varepsilon} B^{j}_{\beta} = g_{hj} B^{j}_{\beta} H_{\alpha}{}^{h}{}_{\varepsilon}. \qquad (4.20)$$

But when we evaluate the covariant derivative of the relation $g_{\beta\varepsilon} = g_{jh} B^{j}_{\beta} B^{h}_{\varepsilon}$ with respect to $u^{\alpha}$, again noting (4.14), (4.15), (4.16), and (4.18), we obtain

$$0 = g_{jh} H_{\alpha}{}^{j}{}_{\beta} B^{h}_{\varepsilon} + g_{jh} B^{j}_{\beta} H_{\alpha}{}^{h}{}_{\varepsilon},$$

which, when taken in conjunction with (4.20), yields

$$g_{hj} B_\varepsilon^h H_\alpha{}^j{}_\beta = 0. \tag{4.21}$$

In accordance with (4.11) it is therefore inferred that *the $H_\alpha{}^j{}_\beta$, regarded as* $\frac{1}{2}m(m+1)$ *contravariant vectors in $T_n(P)$, are normal to $V_m$ at $P$*. This simple conclusion is of paramount importance to all of our subsequent analysis.

Let us substitute from (4.16) in (4.21), after which (4.5) is applied, which gives

$$g_{\varepsilon\lambda} \gamma_\alpha{}^\lambda{}_\beta = g_{hj} B_\varepsilon^h B_\alpha{}^j{}_\beta + g_{lj} B_\varepsilon^l \gamma_h{}^j{}_k B_\alpha^h B_\beta^k,$$

or

$$\gamma_{\alpha\varepsilon\beta} = g_{hj} B_\varepsilon^h B_\alpha{}^j{}_\beta + \gamma_{hjk} B_\alpha^h B_\varepsilon^j B_\beta^k. \tag{4.22}$$

This represents an explicit expression for the Christoffel symbols of $V_m$ in terms of the Christoffel symbols of $V_n$. The formula (4.22) may also be verified (although somewhat laboriously) by direct calculation upon differentiation of (4.5) with respect to $u^\gamma$, and substitution of the relations thus obtained in the explicit definition (4.13).

As an immediate application of (4.21) let us consider a differentiable curve $C : u^\alpha = u^\alpha(t)$ of $V_m$, referred to an arbitrary parameter $t$. The parametric representation of $C$ as a curve of $V_n$ is given by $x^j = x^j(u^\alpha(t))$ in terms of (4.1), so that

$$\dot{x}^j = B_\alpha^j \dot{u}^\alpha, \tag{4.23}$$

where $\dot{x}^j = dx^j/dt$, $\dot{u}^\alpha = du^\alpha/dt$. Let $U^\alpha = U^\alpha(t)$ be a differentiable contravariant vector field of $V_m$ defined along $C$, whose components relative to $V_n$ are given by

$$X^j = B_\alpha^j U^\alpha. \tag{4.24}$$

In terms of (4.17) and (4.16) we then have along $C$:

$$\frac{dX^j}{dt} = B_\alpha{}^j{}_\beta U^\alpha \dot{u}^\beta + B_\alpha^j \frac{dU^\alpha}{dt}$$

$$= H_\alpha{}^j{}_\beta U^\alpha \dot{u}^\beta - \gamma_h{}^j{}_k B_\alpha^h B_\beta^k U^\alpha \dot{u}^\beta + B_\lambda^j \gamma_\alpha{}^\lambda{}_\beta U^\alpha \dot{u}^\beta + B_\alpha^j \frac{dU^\alpha}{dt}$$

$$= H_\alpha{}^j{}_\beta U^\alpha \dot{u}^\beta - \gamma_h{}^j{}_k X^h \dot{x}^k + B_\alpha^j \left( \frac{dU^\alpha}{dt} + \gamma_\varepsilon{}^\alpha{}_\beta U^\varepsilon \dot{u}^\beta \right),$$

where, in the last step, we have used (4.23) and (4.24). From the definition of absolute derivative it therefore follows that

$$\frac{DX^j}{Dt} = H_\alpha{}^j{}_\beta U^\alpha \dot{u}^\beta + B_\alpha^j \frac{DU^\alpha}{Dt}. \tag{4.25}$$

This formula clearly indicates how *the absolute derivative with respect to the embedding space $V_n$ of a tangential vector field $X^j$ is decomposed into normal and tangential components*. Because of (4.21) the first term on the right-hand side of (4.25) is the normal component, while the tangential component is determined by the absolute derivative of the given vector field with respect to the subspace $V_m$. This conclusion is tantamount to the statement that the absolute derivative with respect to the subspace $V_m$ of any vector field tangential to $V_m$ is obtained by the projection onto $V_m$ of its absolute derivative with respect to $V_n$.

In passing we note that this property of the absolute derivative with respect to $V_m$ could actually have been used to define such derivatives; in other words, one could have begun with a decomposition of $DX^j/Dt$ into normal and tangential components, and then have defined $DU^\alpha/Dt$ by the identification of $B_\alpha^j DU^\alpha/Dt$ with the tangential components. This procedure does, in fact, lead to a unique connection on $V_m$, the so-called *induced connection*. Not surprisingly, however, it may be shown that the coefficients of the induced connection are identical with the Christoffel symbols (4.13). The latter are said to determine an *intrinsic connection* on $V_m$, since they are defined in the usual manner in terms of the metric tensor of $V_m$. This state of affairs may be described by the assertion that, *for any subspace of a Riemannian manifold, the induced and intrinsic connections are identical*. (In more general geometries, such as Finsler spaces, this conclusion is not valid.)

As an immediate consequence of (4.25) one may state the following result: *If the vector field $U^\alpha$ is parallel along $C$ (that is, if $DU^\alpha/Dt = 0$ along $C$), then its absolute derivative with respect to $V_n$ is always normal to $V_m$.*

A particularly interesting special case of (4.25) occurs when we identify the vector field $U^\alpha$ along $C$ with the tangent vector $\dot{u}^\alpha$ of $C$, after which we put $t = s$, where $s$ denotes the arc length of $C$ (it being assumed that nowhere along $C$ does $\dot{u}^\alpha$ represent a null direction). Under these circumstances we obtain

$$\frac{Dx'^j}{Ds} = H_\alpha{}^j{}_\beta u'^\alpha u'^\beta + B_\alpha^j \frac{Du'^\alpha}{Ds}. \tag{4.26}$$

Here it should be observed that the first term on the right-hand side depends, at each point $P(u^\alpha)$ of $C$, only on the coordinates $u^\alpha$ of $P$ and the components $u'^\alpha$ of the unit tangent vector of $C$ at $P$. This term is therefore identical for all curves of $V_m$ which pass through $P$ and have the common tangent $u'^\alpha$. However, there exists a unique geodesic $\Gamma$ of $V_m$ which passes through $P$ and possesses the tangent $u'^\alpha$: since $Du'^\alpha/Ds = 0$ for $\Gamma$, it follows from (4.26) that

$$\left( \frac{Dx'^j}{Ds} \right)_\Gamma = H_\alpha{}^j{}_\beta u'^\alpha u'^\beta. \tag{4.27}$$

This result clearly exhibits the geometrical significance of the term in question; incidentally, with the aid of (4.21) we may also conclude that *the principal normal of any geodesic of $V_m$, regarded as a curve of $V_n$, is always normal to the subspace $V_m$.* Also, according to the definition (2.8) of the curvature $1/\rho$ of curves of $V_n$, the curvature $1/\rho_\Gamma$ of $\Gamma$ is given by

$$\left(\frac{1}{\rho_\Gamma}\right)^2 = g_{jh}\left(\frac{Dx'^j}{Ds}\right)_\Gamma \left(\frac{Dx'^h}{Ds}\right)_\Gamma = g_{jh}(H_\alpha{}^j{}_\beta u'^\alpha u'^\beta)(H_\varepsilon{}^h{}_\gamma u'^\varepsilon u'^\gamma). \qquad (4.28)$$

Again, this quantity depends solely on $P$ and the direction $u'^\alpha$ at $P$ and is therefore completely determined at each point and for any tangential direction at that point by the given subspace $V_m$. (In fact, this quantity corresponds to the square of what is generally called the normal curvature in the classical theory of surfaces.)

Returning to the case of an arbitrary differentiable curve $C$ on $V_m$ we now define, by analogy with (2.8), the so-called *geodesic curvature* $1/\rho_g$ by

$$\left(\frac{1}{\rho_g}\right)^2 = g_{\alpha\beta}\frac{Du'^\alpha}{Ds}\frac{Du'^\beta}{Ds}, \qquad (4.29)$$

which is simply the square of the curvature of $C$ when the latter is regarded as a curve of $V_m$. From (4.26), as applied to $C$, we then have

$$g_{jh}\frac{Dx'^j}{Ds}\frac{Dx'^h}{Ds} = g_{jh}(H_\alpha{}^j{}_\beta u'^\alpha u'^\beta)(H_\varepsilon{}^h{}_\gamma u'^\varepsilon u'^\gamma)$$

$$+ 2g_{jh}(H_\alpha{}^j{}_\beta u'^\alpha u'^\beta)B_\varepsilon^h\frac{Du'^\varepsilon}{Ds} + g_{\alpha\beta}\frac{Du'^\alpha}{Ds}\frac{Du'^\beta}{Ds},$$

or, if we invoke (2.8), (4.21), (4.28), and (4.29),

$$\frac{1}{\rho^2} = \left(\frac{1}{\rho_\Gamma}\right)^2 + \left(\frac{1}{\rho_g}\right)^2. \qquad (4.30)$$

This relation clearly displays the relationship between the two curvatures $1/\rho$ and $1/\rho_g$ of any curve $C$ on $V_m$. In particular, if the metric of $V_n$ is positive definite, the right-hand side of (4.28) is always nonnegative, so that in this case

$$\frac{1}{\rho^2} \geq \left(\frac{1}{\rho_g}\right)^2. \qquad (4.31)$$

Let us now turn to the relationship between the respective curvature tensors of $V_n$ and $V_m$. For an $m$-dimensional subspace $X_m$ of a general $n$-dimensional differentiable manifold $X_n$, both $X_m$ and $X_n$ being endowed with arbitrary connections, this relationship is given by (5.7.36). In the present context the respective connections are defined by the Christoffel symbols

of $V_m$ and $V_n$, respectively, for which the torsion tensors are identically zero, so that (5.7.36) reduces to

$$H_\alpha{}^j{}_{\beta\|\gamma} - H_\alpha{}^j{}_{\gamma\|\beta} = R_l{}^j{}_{hk}B_\alpha^l B_\beta^h B_\gamma^k - B_\lambda^j R_\alpha{}^\lambda{}_{\beta\gamma}, \qquad (4.32)$$

in which $R_i{}^j{}_{hk}$, $R_\alpha{}^\lambda{}_{\beta\gamma}$ respectively denote the curvature tensors of $V_n$ and $V_m$, the latter being defined as usual by

$$R_\alpha{}^\lambda{}_{\beta\gamma} = \frac{\partial\gamma_\alpha{}^\lambda{}_\beta}{\partial u^\gamma} - \frac{\partial\gamma_\alpha{}^\lambda{}_\gamma}{\partial u^\beta} + \gamma_\sigma{}^\lambda{}_\gamma\gamma_\alpha{}^\sigma{}_\beta - \gamma_\sigma{}^\lambda{}_\beta\gamma_\alpha{}^\sigma{}_\gamma. \qquad (4.33)$$

The relations (4.32) may be written in a more illuminating way as follows. Let us differentiate (4.21) covariantly with respect to $u^\gamma$, noting (4.16) at the same time, thus obtaining

$$g_{hj}B_\varepsilon^h H_\alpha{}^j{}_{\beta\|\gamma} + g_{hj}H_\varepsilon{}^h{}_\gamma H_\alpha{}^j{}_\beta = 0. \qquad (4.34)$$

This suggests that we multiply (4.32) by $g_{hj}B_\varepsilon^h$; in doing so, we also observe (3.5) and (4.5), which yields

$$g_{hj}B_\varepsilon^h H_\alpha{}^j{}_{\beta\|\gamma} - g_{hj}B_\varepsilon^h H_\alpha{}^j{}_{\gamma\|\beta} = R_{ljhk}B_\alpha^l B_\varepsilon^j B_\beta^h B_\gamma^k - R_{\alpha\varepsilon\beta\gamma}, \qquad (4.35)$$

where, by analogy with (3.5), we have put

$$R_{\alpha\varepsilon\beta\gamma} = g_{\varepsilon\lambda}R_\alpha{}^\lambda{}_{\beta\gamma}. \qquad (4.36)$$

We now substitute from (4.34) in (4.35), which gives

$$-g_{hj}H_\varepsilon{}^h{}_\gamma H_\alpha{}^j{}_\beta + g_{hj}H_\varepsilon{}^h{}_\beta H_\alpha{}^j{}_\gamma = R_{ljhk}B_\alpha^l B_\varepsilon^j B_\beta^h B_\gamma^k - R_{\alpha\varepsilon\beta\gamma},$$

or, on interchange of the indices $\varepsilon$, $\beta$, and $\gamma$

$$R_{\alpha\beta\gamma\varepsilon} = R_{ljhk}B_\alpha^l B_\beta^j B_\gamma^h B_\varepsilon^k + g_{jh}(H_\alpha{}^j{}_\gamma H_\beta{}^h{}_\varepsilon - H_\alpha{}^j{}_\varepsilon H_\beta{}^h{}_\gamma). \qquad (4.37)$$

This is the relation which we have been seeking: it clearly displays *the curvature tensor of the subspace $V_m$ in terms of the curvature tensor of the embedding space $V_n$ and the tensor $H_\alpha{}^j{}_\beta$.* The relation (4.37) is called the *equation of Gauss* since it is the generalization of a formula of the classical differential geometry of surfaces which bears this name.

## 7.5 HYPERSURFACES OF A RIEMANNIAN MANIFOLD

The theory of subspaces assumes a particularly simple and geometrically illuminating form for the case of *hypersurfaces*, namely, when the dimension $m$ of the subspace is $m = n - 1$. This simplification is largely due to the fact that under these circumstances the rank of the matrix $(B_\alpha^j)$ is $n - 1$, which allows us to define a unique unit normal at each point $P$ of $V_{n-1}$. This follows directly from the fact that the system (4.11) possesses a unique solution (up to a multiplicative factor). This solution is normalized, and accordingly

the unit normal $N^j$ to $V_{n-1}$ at $P$ is determined (except for its sign) by the conditions

$$g_{hj}B_\alpha^h N^j = 0, \qquad (5.1)$$

$$\varepsilon(N)g_{hj}N^h N^j = 1, \qquad (5.2)$$

where $\varepsilon(N)$ is the indicator of $N^j$. [Again we exclude from our considerations hypersurfaces possessing points for which the determinant of $g_{\alpha\beta}$ vanishes identically.] We suppose, moreover, that the sign of $N^j$ is chosen such that $N_j = g_{jh}N^h = \phi v_j$, with $\phi > 0$, where the $v_j$ denote the components of a relative covariant vector defined by

$$v_j = (-1)^{j+1}\frac{\partial(x^1,\ldots,x^{j-1},x^{j+1},\ldots,x^n)}{\partial(u^1,\qquad\ldots\qquad,u^{n-1})}, \qquad (5.3)$$

the latter being normal to $V_{n-1}$ by virtue of the identity $v_j B_\alpha^j = 0$.

Having thus defined a unique normal $N^j$ at each point $P$ of $V_{n-1}$, let us return to the equations (4.16) and (4.21). According to the latter the $\frac{1}{2}n(n-1)$ vectors $H_{\alpha\ \beta}^{\ j}$ are normal to $V_{n-1}$; hence we may write

$$B_{\alpha\|\beta}^j = H_{\alpha\ \beta}^{\ j} = \varepsilon(N)N^j\Omega_{\alpha\beta}, \qquad (5.4)$$

where the coefficients $\Omega_{\alpha\beta}$ are class $C^1$ functions of the parameters $u^\varepsilon$ and represent the components of a type $(0,2)$ tensor relative to parameter transformations on $V_{n-1}$. Moreover, because of (4.18), the $\Omega_{\alpha\beta}$ are symmetric. [The quantities thus defined should not be confused with the curvature 2-forms (5.6.39).] The coefficients $\Omega_{\alpha\beta}$ are of basic importance to all subsequent developments; their geometrical significance emerges clearly from the following considerations.

Let $P(u^\alpha)$ be a point on $V_{n-1}$, and let $C$ be some differentiable curve of $V_{n-1}$ passing through $P$ with given unit tangent $u'^\alpha = du^\alpha/ds$ at $P$. According to (2.9) the principal normal $n^j$ of $C$, regarded as a curve of $V_n$, is given by $Dx'^j/Ds = n^j/\rho$, where $1/\rho$ denotes the curvature of $C$. From (4.26) and (5.4) we then have

$$\frac{Dx'^j}{Ds} = \varepsilon(N)N^j\Omega_{\alpha\beta}u'^\alpha u'^\beta + B_\alpha^j\frac{Du'^\alpha}{Ds} = \frac{n^j}{\rho}. \qquad (5.5)$$

In particular, for the unique geodesic $\Gamma$ of $V_{n-1}$ which possesses the tangent $u'^\alpha$ at $P$,

$$\left(\frac{Dx'^j}{Ds}\right)_\Gamma = \varepsilon(N)N^j\Omega_{\alpha\beta}u'^\alpha u'^\beta = \frac{n_\Gamma^j}{\rho_\Gamma}, \qquad (5.6)$$

where $n_\Gamma^j$ denotes the principal normal of $\Gamma$, the latter being regarded as a curve of $V_n$. Thus, since $N^j$, $n_\Gamma^j$ are unit vectors, the curvature $1/\rho_\Gamma$ of $\Gamma$ is

given by

$$\pm\frac{1}{\rho_\Gamma} = \Omega_{\alpha\beta}u'^\alpha u'^\beta = \frac{\Omega_{\alpha\beta}\,du^\alpha\,du^\beta}{g_{\alpha\beta}\,du^\alpha\,du^\beta}. \qquad (5.7)$$

The numerator on the right-hand side of (5.7) is called the *second fundamental form* of the hypersurface $V_{n-1}$; it is a direct generalization of the corresponding concept of the classical theory of surfaces. It is evident from (5.7) that the second fundamental form provides some indication of the curvature properties of $V_{n-1}$ at $P$ in the direction $u'^\alpha$.

Returning to the curve $C$, let us multiply (5.5) by $g_{hj}N^h$, noting (5.1), (5.2), together with the definition (1.38). It is thus found that

$$\Omega_{\alpha\beta}u'^\alpha u'^\beta = \frac{\cos(N,n)}{\rho}. \qquad (5.8)$$

This result implies the well-known *theorem of Meusnier*: since the left-hand side depends solely on the coordinates of the point $P$ and the direction $u'^\alpha$ at $P$, the relation (5.8) implies that *the ratio* $\cos(N,n)/\rho$ *has the same value for all curves of* $V_{n-1}$ *which pass through* $P$ *and possess a common tangent* $u'^\alpha$ *at* $P$. This quantity is called the *normal curvature* $K(u^\varepsilon, u'^\varepsilon)$ of $V_{n-1}$ at $P$ in the direction $u'^\varepsilon$: from (5.7) and (5.8) it follows that, apart from sign, the normal curvature $K(u^\varepsilon, u'^\varepsilon)$ may be identified with the curvature $1/\rho_\Gamma$ of the geodesic $\Gamma$ of $V_{n-1}$ which passes through $P$ in the direction $u'^\alpha$. For the unit vector $u'^\alpha$ the normal curvature is simply the quadratic form $\Omega_{\alpha\beta}u'^\alpha u'^\beta$; for an arbitrary vector $\dot{u}^\alpha$ tangent to $V_{n-1}$ at $P$ one has to write

$$K(u^\varepsilon, \dot{u}^\varepsilon) = \frac{\Omega_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta}{g_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta} = \frac{\Omega_{\alpha\beta}\,du^\alpha\,du^\beta}{g_{\alpha\beta}\,du^\alpha\,du^\beta}. \qquad (5.9)$$

For any given $V_{n-1}$ the coefficients of the first and second fundamental forms can always be evaluated explicitly from the embedding equations (4.1), and hence the normal curvature can be calculated at any point $P$ of $V_{n-1}$ for any prescribed direction $u'^\alpha$ through $P$. Accordingly Meusnier's theorem (5.8) specifies the distribution of the curvatures of the set of all curves of $V_{n-1}$ which pass through $P$ in the direction $u'^\alpha$.

However, it is clearly necessary to consider also curves of $V_{n-1}$ which pass through $P$ in different directions, and therefore one is naturally interested in the values of the normal curvature corresponding to all directions tangential to $V_{n-1}$ at $P$: in particular, those directions at $P$ for which the normal curvature assumes extreme values are of paramount importance. Before proceeding with this program it should be pointed out that the normal curvature, as given by (5.9), is defined solely for non-null directions $\dot{u}^\alpha$. Therefore, in order to avoid undue complications we assume, for the moment,

that there are no null directions in $T_m(P)$, so that, without loss of generality, we may suppose that the first fundamental form $g_{\alpha\beta}\, du^\alpha\, du^\beta$ is positive definite at $P$. The problem of finding extreme values of (5.9) is obviously tantamount to the determination of extreme values of $\Omega_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta$ subject to the restriction $g_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta = 1$. According to the method of Lagrange multipliers we therefore construct the function

$$F(u^\varepsilon, \dot{u}^\varepsilon) = \Omega_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta - \lambda(g_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta - 1),$$

it being recalled that for the required extremum it is necessary that $\partial F/\partial\dot{u}^\gamma = 0$. Because of the symmetry of $\Omega_{\alpha\beta}$ and $g_{\alpha\beta}$ the latter condition is obviously equivalent to

$$(\Omega_{\alpha\gamma} - \lambda g_{\alpha\gamma})\dot{u}^\alpha = 0. \tag{5.10}$$

This is a system of $n-1$ homogeneous linear equations, nontrivial solutions of which correspond to values of $\lambda$ which satisfy the so-called *characteristic equation*

$$\det(\Omega_{\alpha\gamma} - \lambda g_{\alpha\gamma}) = 0. \tag{5.11}$$

The left-hand side of (5.11) is a polynomial of order $n-1$ in $\lambda$; thus (5.11) possesses $n-1$ solutions $\lambda_{(r)}$ $(r = 1, \ldots, n-1)$, which are not necessarily distinct. To each solution $\lambda_{(r)}$ (an eigenvalue) there corresponds a direction $\dot{u}^\alpha_{(r)}$ (the corresponding eigenvector) such that (5.10) is satisfied:

$$\Omega_{\alpha\gamma}\dot{u}^\alpha_{(r)} = \lambda_{(r)}g_{\alpha\gamma}\dot{u}^\alpha_{(r)} \qquad (r = 1, \ldots, n-1). \tag{5.12}$$

When this relation is multiplied by $\dot{u}^\gamma_{(r)}$ (no summation over the index $r$), it follows immediately from (5.9) that $\lambda_{(r)}$ is given by

$$\lambda_{(r)} = K(u^\varepsilon, \dot{u}^\varepsilon_{(r)}). \tag{5.13}$$

These values of the normal curvature are called the *principal curvatures* of $V_{n-1}$ at the point $P(u^\varepsilon)$.

The directions $\dot{u}^\alpha_{(r)}$ satisfying (5.12) are called the *principal directions* of $V_{n-1}$ at $P$; if the $n-1$ roots of (5.11) are all distinct, these directions are uniquely defined. Let $\dot{u}^\alpha_{(r)}$, $\dot{u}^\alpha_{(s)}$ represent two principal directions corresponding to distinct roots $\lambda_{(r)}$, $\lambda_{(s)}$ of (5.11), so that

$$\Omega_{\alpha\gamma}\dot{u}^\alpha_{(r)} = \lambda_{(r)}g_{\alpha\gamma}\dot{u}^\alpha_{(r)}, \quad \Omega_{\alpha\gamma}\dot{u}^\alpha_{(s)} = \lambda_{(s)}g_{\alpha\gamma}\dot{u}^\alpha_{(s)}.$$

Multiplying the first of these by $\dot{u}^\gamma_{(s)}$, the second by $\dot{u}^\gamma_{(r)}$, and subtracting the relations thus obtained, it is found that

$$g_{\alpha\beta}\dot{u}^\alpha_{(r)}\dot{u}^\beta_{(s)} = 0. \tag{5.14}$$

*Thus principal directions corresponding to distinct roots of the characteristic equation* (5.11) *are mutually orthogonal.* Moreover, it is known from the

theory of quadratic forms that one can always construct a set of $n-1$ mutually orthogonal principal directions, even when the roots of (5.11) are not all distinct (in which case, however, the principal directions are not all unique). These directions may be used to construct an orthonormal basis in $T_m(P)$, relative to which the second fundamental form assumes the following diagonal form:

$$\Omega_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta = \lambda_{(1)}(\dot{u}^1)^2 + \lambda_{(2)}(\dot{u}^2)^2 + \cdots + \lambda_{(n-1)}(\dot{u}^{n-1})^2 = K(u^\varepsilon, \dot{u}^\varepsilon)g_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta. \tag{5.15}$$

This particular representation of normal curvature in terms of the principal curvatures (5.13) is the counterpart of the *theorem of Euler* for two-dimensional surfaces $V_2$ embedded in a three-dimensional Euclidean space.

Let us put

$$\Omega^\alpha_\beta = g^{\alpha\gamma}\Omega_{\beta\gamma}, \tag{5.16}$$

noting at the same time that the characteristic equation (5.11) may then be written in the equivalent form

$$\det(\Omega^\alpha_\beta - \lambda\delta^\alpha_\beta) = 0, \tag{5.17}$$

to which the theory of Section 4.2 may be applied directly. *The fundamental invariants of $V_{n-1}$ are defined to be the $n-1$ elementary symmetric functions of the roots* $\lambda_{(1)}, \ldots, \lambda_{(n-1)}$ *of* (5.17):

$$\begin{aligned}
H_{(1)} &= \lambda_{(1)} + \cdots + \lambda_{(n-1)}, \\
H_{(2)} &= \lambda_{(1)}\lambda_{(2)} + \lambda_{(1)}\lambda_{(3)} + \cdots + \lambda_{(n-2)}\lambda_{(n-1)}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
H_{(n-1)} &= \lambda_{(1)}\lambda_{(2)}\cdots\lambda_{(n-1)}.
\end{aligned} \tag{5.18}$$

The formula (4.2.43) allows us to compute these invariants directly as polynomials in $\Omega^\alpha_\beta$; in particular, for the so-called *mean curvature* $H_{(1)}$ we have

$$H_{(1)} = \Omega^\alpha_\alpha = g^{\alpha\beta}\Omega_{\alpha\beta},$$

while

$$H_{(n-1)} = (-1)^{n-1}\det(\Omega^\alpha_\beta) = (-1)^{n-1}\det(g^{\alpha\gamma}\Omega_{\beta\gamma}) = (-1)^{n-1}\frac{\det(\Omega_{\alpha\beta})}{\det(g_{\alpha\beta})}. \tag{5.19}$$

When $n = 3$, the latter expression is known as the *Gaussian curvature* of the surface $V_2$.

A further very significant property of the coefficients of the second fundamental form is the fact that the latter allow us to express the covariant derivative of the unit normal $N^j$ in a very elegant manner. Differentiating

(5.1) with respect to $u^\beta$, noting (4.15), (5.4), and (5.2), we obtain

$$g_{hj}B_\alpha^h N_{\parallel\beta}^j = -g_{hj}N^j B_{\alpha\parallel\beta}^h = -g_{hj}N^h N^j \varepsilon(N)\Omega_{\alpha\beta} = -\Omega_{\alpha\beta}, \qquad (5.20)$$

while differentiation of (5.2) yields

$$g_{hj}N^h N_{\parallel\beta}^j = 0, \qquad (5.21)$$

which implies that $N_{\parallel\beta}^j$ is tangential to $V_{n-1}$. Thus there exist coefficients $\lambda_\beta^\varepsilon$ such that $N_{\parallel\beta}^j = B_\varepsilon^j \lambda_\beta^\varepsilon$. When this is substituted in (5.20), the relations (4.5) being taken into account, it is seen that

$$g_{hj}B_\alpha^h N_{\parallel\beta}^j = g_{hj}B_\alpha^h B_\varepsilon^j \lambda_\beta^\varepsilon = g_{\alpha\varepsilon}\lambda_\beta^\varepsilon = -\Omega_{\alpha\beta},$$

so that $\lambda_\beta^\varepsilon = -\Omega_\beta^\varepsilon$ by (5.16). We therefore conclude that *the covariant derivative of the unit normal is given by*

$$N_{\parallel\beta}^j = -B_\varepsilon^j \Omega_\beta^\varepsilon. \qquad (5.22)$$

Thus along an arbitrary curve $C$ of $V_{n-1}$ the absolute derivative of $N^j$ is simply

$$\frac{DN^j}{Ds} = N_{\parallel\beta}^j \frac{du^\beta}{ds} = -B_\varepsilon^j \Omega_\beta^\varepsilon u'^\beta = -B_\varepsilon^j g^{\alpha\varepsilon}\Omega_{\alpha\beta}u'^\beta. \qquad (5.23)$$

A curve of $V_{n-1}$ is said to be a *line of curvature* of $V_{n-1}$ if its tangent vector coincides everywhere with a principal direction. Along such a curve we have, by virtue of (5.10) and (5.13),

$$\Omega_{\alpha\beta}u'^\beta = K(u, u')g_{\alpha\beta}u'^\beta, \qquad (5.24)$$

where $K(u, u')$ denotes the principal curvature in the direction of $u'^\beta$. When (5.24) is substituted in (5.23) the latter reduces to

$$\frac{DN^j}{Ds} = -K(u, u')B_\varepsilon^j g^{\alpha\varepsilon}g_{\alpha\beta}u'^\beta = -K(u, u')B_\beta^j u'^\beta = -K(u, u')x'^j. \qquad (5.25)$$

This simple conclusion implies the *theorem of Rodrigues: along a line of curvature of $V_{n-1}$ the direction of the absolute derivative of the unit normal of $V_{n-1}$ coincides with the direction of the tangent of the line of curvature.*

If the second fundamental form is not positive definite at a point $P$ of $V_{n-1}$, there exist directions $u'^\alpha$ at $P$ for which

$$\Omega_{\alpha\beta}u'^\alpha u'^\beta = 0. \qquad (5.26)$$

These are called *asymptotic directions*, and a curve of $V_{n-1}$ whose tangent vectors coincide everywhere with such directions is called an *asymptotic line* of $V_{n-1}$. For curves of this kind the relation (5.5) reduces to

$$\frac{Dx'^j}{Ds} = B_\alpha^j \frac{Du'^\alpha}{Ds} = \frac{n^j}{\rho}, \qquad (5.27)$$

which indicates that the principal normal of an asymptotic line, regarded as a curve of $V_n$, is always tangential to $V_{n-1}$. If, in particular, the hypersurface $V_{n-1}$ is such that it contains asymptotic lines which are also geodesics of $V_{n-1}$, it follows from (5.27) that $Dx'^j/Ds = 0$, that is, such curves are also geodesics of $V_n$. Conversely, if the hypersurface $V_{n-1}$ contains a geodesic of $V_n$, this curve is an asymptotic line as well as a geodesic of $V_{n-1}$. It should be emphasized, however, that whereas principal directions exist at each point $P$ of any $V_{n-1}$, this is not true for asymptotic directions. In fact, if all the roots $\lambda_{(r)}$ of the characteristic equation (5.11) are positive, it is evident from (5.15) that condition (5.26) cannot be satisfied.

Let us now turn to the relationship between the respective curvature tensors of $V_n$ and $V_{n-1}$. First, if we substitute from (5.4) in (4.37), noting (5.2), we obtain

$$R_{\alpha\beta\gamma\varepsilon} = R_{ljhk}B_\alpha^l B_\beta^j B_\gamma^h B_\varepsilon^k + \varepsilon(N)(\Omega_{\alpha\gamma}\Omega_{\beta\varepsilon} - \Omega_{\alpha\varepsilon}\Omega_{\beta\gamma}). \qquad (5.28)$$

This is the *equation of Gauss for a hypersurface $V_{n-1}$ of an n-dimensional Riemannian manifold $V_n$*; it is without doubt the most significant relation in the theory of hypersurfaces. *The first term on the right-hand side of (5.28) clearly displays the influence exerted by the curvature tensor of the embedding space $V_n$ on the curvature of the embedded manifold $V_{n-1}$, while the second term gives a precise portrayal of the contribution due to the coefficients of the second fundamental form of $V_{n-1}$.*

Second, let us differentiate (5.4) with respect to $u^\gamma$, at the same time observing (5.22), which yields

$$H_{\alpha\ \beta\parallel\gamma}^{\ j} = \varepsilon(N)(N^j \Omega_{\alpha\beta\parallel\gamma} - B_\varepsilon^j \Omega_\gamma^\varepsilon \Omega_{\alpha\beta}). \qquad (5.29)$$

When this is multiplied by $N_j = g_{hj}N^h$, the relations (5.1) and (5.2) being taken into account, it is found that

$$N_j H_{\alpha\ \beta\parallel\gamma}^{\ j} = \Omega_{\alpha\beta\parallel\gamma}. \qquad (5.30)$$

But from (4.32) we have, again because of (5.1),

$$N_j(H_{\alpha\ \beta\parallel\gamma}^{\ j} - H_{\alpha\ \gamma\parallel\beta}^{\ j}) = N_j R_{i\ hk}^{\ j} B_\alpha^l B_\beta^h B_\gamma^k, \qquad (5.31)$$

and hence by (5.30),

$$\Omega_{\alpha\beta\parallel\gamma} - \Omega_{\alpha\gamma\parallel\beta} = N_j R_{i\ hk}^{\ j} B_\alpha^l B_\beta^h B_\gamma^k. \qquad (5.32)$$

These are the *equations of Codazzi for a hypersurface $V_{n-1}$ of $V_n$*. They are essentially integrability conditions on $\Omega_{\alpha\beta}$ and indicate quite clearly that the components of an arbitrary symmetric type (0, 2) tensor field cannot, in general, play the role of the coefficients of the second fundamental form of some hypersurface $V_{n-1}$.

Let us conclude this section with a few remarks of a general nature concerning the case when the embedding manifold $V_n$ is a Euclidean space $E_n$. Under these circumstances $R_l{}^j{}_{hk} = 0$, and the equations (5.28) and (5.32) of Gauss and Codazzi reduce to

$$R_{\alpha\beta\gamma\varepsilon} = \Omega_{\alpha\gamma}\Omega_{\beta\varepsilon} - \Omega_{\alpha\varepsilon}\Omega_{\beta\gamma}, \tag{5.33}$$

and

$$\Omega_{\alpha\beta\|\gamma} - \Omega_{\alpha\gamma\|\beta} = 0, \tag{5.34}$$

respectively. Irrespective of these simplifications, the theory of hypersurfaces continues to depend on the fundamental tensors $g_{\alpha\beta}$ and $\Omega_{\alpha\beta}$. The existence of the former renders any hypersurface $V_{n-1}$ of $E_n$ a Riemannian space in its own right; any analytical development which depends solely on $g_{\alpha\beta}$ and the derivatives thereof is part of the *intrinsic theory* of $V_{n-1}$ in the sense that, once the $g_{\alpha\beta}$ are given, no appeal whatsoever need be made to the fact that $V_{n-1}$ is embedded in $E_n$. On the other hand, *the coefficients $\Omega_{\alpha\beta}$ of the second fundamental form are defined if and only if an embedding is assumed*: therefore, a development which includes the use of these quantities cannot be expected to be necessarily intrinsic. Indeed, it is obvious that any statement pertaining to the normals of $V_{n-1}$ must be nonintrinsic; the same applies to concepts such as principal or asymptotic directions. However, when we consider the equation (5.33) of Gauss in this context, we are confronted with the following remarkable phenomenon. The left-hand side of (5.33), namely, the curvature tensor of $V_{n-1}$, is clearly intrinsic, since this tensor is defined entirely in terms of the $g_{\alpha\beta}$ together with their first and second derivatives. The right-hand side, however, is merely a bilinear form in the coefficients $\Omega_{\alpha\beta}$, and therefore consists of elements each of which is nonintrinsic. One is therefore led to conclude that there must exist intrinsic functions of the nonintrinsic coefficients, and that the right-hand side of (5.33) is an example of a function of this kind.

The significance of this example may be illustrated most effectively for the case when $n = 3$. From (5.19) it is evident that the Gaussian curvature of a surface $V_2$ is given by

$$H_{(2)} = \frac{\det(\Omega_{\alpha\beta})}{\det(g_{\alpha\beta})}. \tag{5.35}$$

Moreover, from the theory of Section 7.3 we recall that, for any $V_2$, all nonvanishing components of the curvature tensor are of the form $\pm R_{1212}$. Thus only a single independent Gauss equation can be obtained from (5.33), namely, when we put $\alpha = \gamma = 1$, $\beta = \varepsilon = 2$, which gives

$$R_{1212} = \Omega_{11}\Omega_{22} - \Omega_{12}\Omega_{21} = \det(\Omega_{\alpha\beta}). \tag{5.36}$$

The expression (5.35) for the Gaussian curvature of a $V_2$ embedded in $E_3$ thus assumes the form

$$H_{(2)} = \frac{R_{1212}}{\det(g_{\alpha\beta})}. \tag{5.37}$$

It therefore follows that *the Gaussian curvature is an intrinsic invariant, despite the nonintrinsic appearance of the definition* (5.35). This conclusion, first discovered by Gauss [1] on the basis of some exceedingly intricate calculations, was designated by him the *Theorema Egregium*.

### 7.6  THE DIVERGENCE THEOREM FOR HYPERSURFACES OF A RIEMANNIAN MANIFOLD

In Section 5.5 a divergence theorem for closed hypersurfaces of a differentiable manifold $X_n$ was presented. This theorem is entirely independent of the notion of a metric; however, when the embedding space is Riemannian, in which case a metric is available, a corresponding divergence theorem may be derived in which the role played by the metric is clearly displayed. We shall now address ourselves to this matter.

To this end we shall obtain an identity which relates the determinants of the respective metric tensors of the embedding space $V_n$ and the hypersurface $V_{n-1}$. This identity involves the quantities $v_j$ on $V_{n-1}$ as defined by (5.3), which can be written in an equivalent manner as

$$(n - 1)! v_j = \varepsilon_{jj_2\cdots j_n} J^{j_2\cdots j_n}, \tag{6.1}$$

where

$$J^{j_2\cdots j_n} = \frac{\partial(x^{j_2}, \ldots, x^{j_n})}{\partial(u^1, \ldots, u^{n-1})}. \tag{6.2}$$

Now, as a result of (4.5) we have

$$\begin{aligned}
(n - 1)! \det(g_{\alpha\beta}) &= \varepsilon^{\alpha_1\cdots\alpha_{n-1}}\varepsilon^{\beta_1\cdots\beta_{n-1}}g_{\alpha_1\beta_1}\cdots g_{\alpha_{n-1}\beta_{n-1}} \\
&= \varepsilon^{\alpha_1\cdots\alpha_{n-1}}\varepsilon^{\beta_1\cdots\beta_{n-1}}B_{\alpha_1}^{j_2}\cdots B_{\alpha_{n-1}}^{j_n}B_{\beta_1}^{h_2}\cdots B_{\beta_{n-1}}^{h_n}g_{j_2h_2}\cdots g_{j_nh_n} \\
&= J^{j_2\cdots j_n}J^{h_2\cdots h_n}g_{j_2h_2}\cdots g_{j_nh_n}.
\end{aligned} \tag{6.3}$$

Moreover, by virtue of (4.2.20), (4.2.11), and (4.2.15) the definition (6.1) yields

$$\begin{aligned}
v_j\varepsilon^{jj_2\cdots j_n} &= [(n - 1)!]^{-1}\varepsilon^{jj_2\cdots j_n}\varepsilon_{jh_2\cdots h_n}J^{h_2\cdots h_n} \\
&= [(n - 1)!]^{-1}\delta_{jh_2\cdots h_n}^{jj_2\cdots j_n}J^{h_2\cdots h_n} = [(n - 1)!]^{-1}\delta_{h_2\cdots h_n}^{j_2\cdots j_n}J^{h_2\cdots h_n} = J^{j_2\cdots j_n},
\end{aligned} \tag{6.4}$$

so that (6.3) is equivalent to

$$(n - 1)! \det(g_{\alpha\beta}) = v_j v_h \varepsilon^{jj_2\cdots j_n}\varepsilon^{hh_2\cdots h_n}g_{j_2h_2}\cdots g_{j_nh_n}. \tag{6.5}$$

But according to (4.2.44) the cofactor $G^{jh}$ of $g_{jh}$ in $\det(g_{lk})$ is given by

$$(n-1)! \, G^{jh} = \varepsilon^{jj_2\cdots j_n} \varepsilon^{hh_2\cdots h_n} g_{j_2 h_2} \cdots g_{j_n h_n},$$

and hence, since $G^{jh} = g^{jh}\det(g_{lk})$, the relation (6.5) reduces to

$$\det(g_{\alpha\beta}) = \det(g_{lk}) g^{jh} v_j v_h, \tag{6.6}$$

which is the required identity.

As an immediate consequence of (6.6) we can now obtain an explicit representation of the unit normal $N^j$ of $V_{n-1}$. Let us recall that $v_j B_\alpha^j = 0$ identically; according to the convention stipulated at the beginning of the preceding section we therefore have

$$v_j = v N_j, \tag{6.7}$$

where

$$v = +[\varepsilon(v) g_{hj} v^h v^j]^{1/2}. \tag{6.8}$$

In particular, it follows from (6.6) that *a necessary and sufficient condition that the unit normal $N^j$ at a point $P$ of $V_{n-1}$ define a null-direction in $V_n$ is that the determinant of the induced metric tensor of $V_{n-1}$ vanishes at $P$.*

Since our subsequent analysis will be concerned with closed hypersurfaces in $V_n$, to which the identity (6.6) is to be applied, we assume for the remainder of this section that the metric of $V_n$ is positive definite, thus excluding the possibility of normal null directions.

Now let us turn to the general nonmetric divergence theorem as formulated in equation (5.5.33), in which the integrand of the surface integral involves the $(n-1)$-forms $\pi_j$ as given by (5.5.35). In view of (6.1) and (6.2) the latter can be expressed in the form

$$\pi_j = v_j \, du^1 \wedge \cdots \wedge du^{n-1}.$$

Thus the integral formula (5.5.33) is equivalent to

$$\int_G A^j_{|j} \, dx^1 \wedge \cdots \wedge dx^n = \int_{\partial G} v_j A^j \, du^1 \wedge \cdots \wedge du^{n-1}, \tag{6.9}$$

where $\partial G$ is the hypersurface bounding an $n$-dimensional region $G$ of $V_n$ subject to the restrictions specified in Section 5.5, while $A^j(x^h)$ is a differentiable contravariant relative vector field of weight $+1$ defined on $G$. If $X^j(x^h)$ denotes an arbitrary type $(1, 0)$ absolute tensor field on $G$, then $\sqrt{\det(g_{lk})}\, X^j$ represents the components of a relative tensor field of the same type as $A^j$ since $\sqrt{\det(g_{lk})}$ is a scalar density. We may therefore put $A^j = \sqrt{\det(g_{lk})}\, X^j$ in (6.9), which, with the aid of Ricci's lemma, will then give

$$\int_G \sqrt{\det(g_{lk})}\, X^j_{|j}\, dx^1 \wedge \cdots \wedge dx^n = \int_{\partial G} v_j \sqrt{\det(g_{lk})}\, X^j \, du^1 \wedge \cdots \wedge du^{n-1}. \tag{6.10}$$

In this relation we now substitute from (6.7), noting (6.6) and (6.8) at the same time, thus obtaining

$$\int_G \sqrt{\det(g_{lk})}\, X^j_{|j}\, dx^1 \wedge \cdots \wedge dx^n = \int_{\partial G} \sqrt{\det(g_{\alpha\beta})}\, N_j X^j \, du^1 \wedge \cdots \wedge du^{n-1}. \tag{6.11}$$

*This is the divergence theorem for an arbitrary differentiable contravariant vector field $X^j$ defined over an $n$-dimensional region $G$ of a Riemannian manifold $V_n$.*

The $n$-form

$$\sqrt{\det(g_{lk})}\, dx^1 \wedge \cdots \wedge dx^n \tag{6.12}$$

is interpreted as the $n$-dimensional volume element of $V_n$. This form is a scalar, since $n!(dx^1 \wedge \cdots \wedge dx^n) = \varepsilon_{j_1\cdots j_n} dx^{j_1} \wedge \cdots \wedge dx^{j_n}$ is a relative tensor of weight $-1$. By the same token, the $(n-1)$-form

$$\sqrt{\det(g_{\alpha\beta})}\, du^1 \wedge \cdots \wedge du^{n-1} \tag{6.13}$$

is regarded as the $(n-1)$-dimensional volume element of $V_{n-1}$. From the manner in which the volume elements (6.12) and (6.13) enter into the integral formula (6.11), it should be immediately evident how the latter contains the Gauss divergence theorem of classical vector analysis as a special case.

As an application of the divergence theorem (6.11) we now derive the Gauss-Bonnet theorem for a two-dimensional Riemannian space $V_2$ with positive definite metric (without assuming that $V_2$ is embedded in a manifold of higher dimension). To this end we shall require a certain identity, which will first be derived for a $V_n$ of arbitrary dimension $n$. Let $X^j = X^j(x^h)$ denote a class $C^2$ contravariant vector field defined over a finite region of $V_n$, and let us put

$$-\tfrac{1}{4} Z^j = \mu^{-1} \sqrt{g}\, \delta^{jh}_{lm} X^l X^m_{|h}, \tag{6.14}$$

where

$$\mu = |X|^2 = g_{hj} X^h X^j, \tag{6.15}$$

and

$$g = \det(g_{lk}). \tag{6.16}$$

The divergence of (6.14) is given by

$$-\tfrac{1}{4} Z^j_{|j} = \mu^{-1} \sqrt{g}\, \delta^{jh}_{lm}(X^l X^m_{|h|j} + X^l_{|j} X^m_{|h}) - \mu^{-2} \mu_{|j} \sqrt{g}\, \delta^{jh}_{lm} X^l X^m_{|h}. \tag{6.17}$$

As regards the first term on the right-hand side we note that, by virtue of (3.3), (3.6), and (3.19),

$$\delta^{jh}_{lm} X^l X^m_{|h|j} = \tfrac{1}{2} \delta^{jh}_{lm} X^l (X^m_{|h|j} - X^m_{|j|h})$$
$$= \tfrac{1}{2} \delta^{jh}_{lm} X^l X^k R_k{}^m{}_{hj} = \tfrac{1}{2} X^l X^k (R_k{}^m{}_{ml} - R_k{}^m{}_{lm}) = -R_{lk} X^l X^k. \tag{6.18}$$

For the sake of brevity we now introduce the notation

$$R^{mp}{}_{hk} = g^{mj}R_j{}^p{}_{hk}, \qquad (6.19)$$

observing that, because of (3.10),

$$R^{mp}{}_{hk} = -R^{pm}{}_{hk}. \qquad (6.20)$$

By means of (4.2.9) we then deduce with the aid of (3.3) and (3.19) that

$$
\begin{aligned}
\delta^{jhk}_{lmp} X_j X^l R^{mp}{}_{hk} &= [\delta^j_l(\delta^h_m\delta^k_p - \delta^h_p\delta^k_m) + (\delta^j_m\delta^{hk}_{pl} + \delta^j_p\delta^{hk}_{lm})] X_j X^l R^{mp}{}_{hk} \\
&= \mu(R^{hk}{}_{hk} - R^{kh}{}_{hk}) + X_m X^l \delta^{hk}_{pl} R^{mp}{}_{hk} + X_m X^l \delta^{hk}_{lp} R^{pm}{}_{hk} \\
&= 2\mu R + 4X_m X^l R^{mp}{}_{pl} = 2\mu R - 4R_{lk} X^l X^k,
\end{aligned}
\qquad (6.21)
$$

and thus (6.18) can be written in the form

$$\delta^{jh}_{lm} X^l X^m_{|h|j} = \tfrac{1}{4}\delta^{jhk}_{lmp} X_j X^l R^{mp}{}_{hk} - \tfrac{1}{2}\mu R. \qquad (6.22)$$

An expansion similar to that carried out in (6.21) gives

$$\mu\delta^{jh}_{lm} X^l_{|j} X^m_{|h} = \delta^{jhk}_{lmp} X_j X^l X^m_{|h} X^p_{|k} - 2X_m X^l \delta^{hk}_{pl} X^m_{|h} X^p_{|k}. \qquad (6.23)$$

Also, since

$$\mu_{|j} = 2g_{hl} X^h X^l_{|j},$$

we have

$$
\begin{aligned}
\delta^{jh}_{lm} X^l_{|j} X^m_{|h} - \mu^{-1}\mu_{|j}\delta^{jh}_{lm} X^l X^m_{|h} &= \delta^{hk}_{mp} X^m_{|h} X^p_{|k} + 2\mu^{-1} X_m X^l \delta^{hk}_{pl} X^m_{|h} X^p_{|k} \\
&= \mu^{-1}\delta^{jhk}_{lmp} X_j X^l X^m_{|h} X^p_{|k},
\end{aligned}
\qquad (6.24)
$$

the last step resulting directly from (6.23). We now substitute from (6.22) and (6.24) in (6.17), which yields

$$2\sqrt{g}\,R = Z^j_{|j} + \mu^{-1}\sqrt{g}\,\delta^{jhk}_{lmp} X_j X^l (4\mu^{-1} X^m_{|h} X^p_{|k} + R^{mp}{}_{hk}). \qquad (6.25)$$

This is the identity (Horndeski [1]) which we have been seeking: *it expresses the scalar curvature R of our $V_n$ in terms of the first and second covariant derivatives of an arbitrary class $C^2$ contravariant vector field.*

Let us now restrict ourselves to the case when $n = 2$. Recalling the fact that the generalized Kronecker delta vanishes identically whenever the number of its sub- or superscripts exceeds the dimension, we infer from (6.25) that *the scalar curvature of our $V_2$ may be represented as a divergence*:

$$2\sqrt{g}\,R = Z^j_{|j}. \qquad (6.26)$$

Moreover, for a $V_2$ the scalar curvature $R$ is proportional to the Gaussian curvature $H_{(2)}$ (which, as was shown at the end of the last section is an intrinsic quantity); in fact, from (3.40) and (5.37) we have that

$$R = 2H_{(2)}. \qquad (6.27)$$

In order to apply the divergence theorem to (6.26), let us suppose that the vector field $X^j$ which appears in the definition (6.14) of $Z^j$ is a unit field, that is, that $\mu = 1$, and let us put

$$S^j = -\delta^{jh}_{lm} X^l X^m_{|h} = X^h X^j_{|h} - X^j X^h_{|h}, \qquad (6.28)$$

so that (6.14) becomes

$$Z^j = 4\sqrt{g}\,S^j. \qquad (6.29)$$

Therefore, because of (6.27), we now have

$$H_{(2)} = S^j_{|j}. \qquad (6.30)$$

Now, let $G$ be a simply connected region of $V_2$ bounded by a simple closed curve $C$ of class $C^2$, whose parametric representation is given by $x^j = x^j(s)$, where $s$ denotes the arc length of $C$. Thus $x'^j = dx^j/ds$ is the unit tangent vector of $C$, and the counterpart of (4.5) with $m = 1$ is simply $g_{\alpha\beta} = g_{hj} x'^h x'^j = 1$ (with $\alpha = 1$, $\beta = 1$). Thus if we apply the integral formula (6.11) with $n = 2$ to the vector field $S^j$, we obtain by virtue of (6.30)

$$\int_G \sqrt{g}\,H_{(2)}\,dx^1 \wedge dx^2 = \int_G \sqrt{g}\,S^j_{|j}\,dx^1 \wedge dx^2 = \oint_C N_j\,S^j\,ds, \qquad (6.31)$$

where the unit outward normal $N_j$ is given by

$$N_j = \sqrt{g}\,\varepsilon_{jh} x'^h, \qquad (6.32)$$

in consequence of (6.1), (6.7), and (6.6).

The left-hand side of (6.31) is the so-called *curvatura integra* of the region $G$ of $V_2$, and it remains to reduce the integral on the right-hand side. By means of (6.28), (6.32), and (4.2.20) it is easily verified by direct expansion and use of the fact that $X^j_{|h} x'^h = DX^j/Ds$ along $C$, that

$$N_j S^j = -\sqrt{g}\,\varepsilon_{jh} X^j \frac{DX^h}{Ds}. \qquad (6.33)$$

Since the field $X^j$ is unit by hypothesis, we can decompose $X^j$ into tangential and normal components along $C$ according to

$$X^j = a(s)x'^j + b(s)N^j, \qquad (6.34)$$

with

$$a^2 + b^2 = 1. \qquad (6.35)$$

On taking the covariant derivative of (6.34) with respect to $s$, we obtain

$$
\begin{aligned}
\varepsilon_{jh} X^j \frac{DX^h}{Ds} = \varepsilon_{jh} x'^j N^h(ab' - a'b) &+ a^2\varepsilon_{jh} x'^j \frac{Dx'^h}{Ds} \\
&+ b^2\varepsilon_{jh} N^j \frac{DN^h}{Ds} + ab\varepsilon_{jh}\left(x'^j\frac{DN^h}{Ds} + N^h\frac{Dx'^j}{Ds}\right).
\end{aligned}
\qquad (6.36)
$$

We now reduce each of these expressions in turn. From (6.32) we have

$$\varepsilon_{jh} x'^j N^h = -(\sqrt{g})^{-1} N_h N^h = -(\sqrt{g})^{-1}. \qquad (6.37)$$

Moreover, according to the first Frenet formula $Dx'^j/Ds = n^j/\rho$, where $1/\rho$ is the curvature of $C$, while $n^j$ is the principal normal of $C$, given by $n_1 = -\sqrt{g}\, x'^2,\, n_2 = \sqrt{g}\, x'^1$. Hence by (6.32),

$$\frac{Dx'^j}{Ds} = -\frac{N^j}{\rho}, \qquad (6.38)$$

and thus

$$\varepsilon_{jh} x'^j \frac{Dx'^h}{Ds} = -\rho^{-1} \varepsilon_{jh} x'^j N^h = \rho^{-1} (\sqrt{g})^{-1} N_h N^h = (\rho\sqrt{g})^{-1}. \quad (6.39)$$

Also, since $g_{hj} x'^h N^j = 0$, it may be inferred from (6.32) and (6.38) that

$$\varepsilon_{jh} N^j \frac{DN^h}{Ds} = (\sqrt{g})^{-1} g_{jh} x'^h \frac{DN^j}{Ds} = -(\sqrt{g})^{-1} g_{jh} \frac{Dx'^h}{Ds} N^j = (\rho\sqrt{g})^{-1}, \quad (6.40)$$

while similarly

$$\varepsilon_{jh} x'^j \frac{DN^h}{Ds} = -(\sqrt{g})^{-1} N_h \frac{DN^h}{DS} = 0, \qquad (6.41)$$

together with

$$\varepsilon_{jh} N^j \frac{Dx'^j}{Ds} = -\rho^{-1} \varepsilon_{jh} N^j N^h = 0. \qquad (6.42)$$

On substituting from (6.37) and (6.39)–(6.42) in (6.36), noting (6.35) at the same time, we obtain

$$\varepsilon_{jh} X^j \frac{DX^h}{Ds} = -(\sqrt{g})^{-1}(ab' - a'b) + (\rho\sqrt{g})^{-1}. \qquad (6.43)$$

Because of (6.35) we may put $a = \cos\theta,\, b = \sin\theta$, so that $ab' - a'b = d\theta/ds$, and substitution of (6.43) in (6.33) therefore yields

$$N_j S^j = \frac{d\theta}{ds} - \frac{1}{\rho}. \qquad (6.44)$$

The integral formula (6.31) therefore assumes the form

$$\int_G \sqrt{g}\, H_{(2)}\, dx^1 \wedge dx^2 + \oint_C \frac{ds}{\rho} = 2\pi. \qquad (6.45)$$

This is the famous Gauss-Bonnet theorem for a region $G$ of a two-dimensional Riemannian space $V_2$. In this formulation it is assumed that $G$ is bounded

by a simple closed curve of class $C^2$; it is not difficult to obtain extensions of this theorem when various conditions on $G$ assumed above are relaxed. (Chern [1], Rund [9], Stoker [1]). It should be emphasized that the above derivation of (6.45) is not based on any assumption to the effect that $V_2$ is embedded in a manifold of higher dimension. Naturally (6.45) is valid also if $V_2$ happens to be a subspace; however, in any case it should be clearly understood that the integrand $1/\rho$ refers to the curvature of $C$ as a curve of $V_2$ and is therefore the *geodesic* curvature of $C$ whenever an embedding is assumed.

## PROBLEMS

7.1 By using an appropriate quotient theorem, prove that $g^{ij}$ are the components of a type (2, 0) tensor. (*Hint*: multiply $g^{ij} g_{jk} = \delta^i_k$ by an arbitrary vector $X^k$.) See also Problem 3.23.

7.2 In a three-dimensional Euclidean space the components of the metric tensor referred to orthogonal Cartesian coordinates are simply the entries of the $3 \times 3$ unit matrix, since in these coordinates $ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$. Evaluate the components of the metric tensor referred to a spherical polar coordinate system $(r, \theta, \phi)$.

7.3 In $V_n$, with positive definite metric, let $\Omega$ be a $p$-form $\Omega = a_{i_1 \cdots i_p} dx^{i_1} \wedge \cdots \wedge dx^{i_p}$. An $(n - p)$-form $*\Omega$ is defined by

$$(n - p)!\, {}^*\Omega = \eta_{i_1 \cdots i_n} g^{i_1 j_1} \cdots g^{i_p j_p} a_{j_1 \cdots j_p} dx^{i_{p+1}} \wedge \cdots \wedge dx^{i_n},$$

where

$$\eta_{i_1 \cdots i_n} = \sqrt{g}\, \varepsilon_{i_1 \cdots i_n} \quad \text{and} \quad g = \det(g_{ij}).$$

Show that

$$*(*\Omega) = (-1)^{p(n-p)} \Omega.$$

7.4 In a $V_2$ with line-element

$$ds^2 = \left[ \frac{1}{u^2} - \frac{r^2}{u^4} \right] dr^2 + \frac{r^2}{u^2}\, d\theta^2$$

where $u^2 = r^2 - a^2$ ($a$ constant), $r > a$, show that geodesics are characterized by

$$a^2 \left[ \left( \frac{dr}{d\theta} \right)^2 + r^2 \right] = br^4$$

where $b$ is a constant. To what type of geodesic does $b = 1$ correspond?

7.5 Show that the geodesics of $V_2$ corresponding to the line-element

$$ds^2 = (dx^1)^2 + (x^1)^2 (dx^2)^2$$

are given by $x^1 = a \sec(x^2 + b)$, $a, b$ constants.

7.6    A particular $V_3$ has the line-element

$$ds^2 = dt^2 - \frac{1}{1 - \lambda r^2} dr^2 - r^2 d\phi^2$$

where $\lambda$ is a positive constant and $r^2 < 1/\lambda$. Show that null geodesics are characterized by

$$\left(\frac{dr}{d\phi}\right)^2 = r^2(1 - \lambda r^2)(\alpha r^2 - 1)$$

where $\alpha$ is a constant. By setting $r^2 = 1/u$ solve this differential equation and establish that these geodesics are ellipses if $r, \phi$ denote the usual polar coordinates.

7.7    A particular $V_4$ has the (de-Sitter) line-element

$$ds^2 = F dt^2 - F^{-1} dr^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\phi^2$$

where $F = 1 - r^2/a^2$, ($a$ = constant). Show that, in the plane $\theta = \pi/2$, null geodesics are characterized by

$$\frac{dr}{d\phi} = r\sqrt{\alpha^2 r^2 - 1}$$

where $\alpha$ is a constant, and, furthermore, these are straight lines if $r, \phi$ denote the usual polar coordinates.

7.8    The line-element of a certain $V_4$ is

$$ds^2 = (dx)^2 + (dy)^2 + (dz)^2 + 2gt \, dx \, dt - c^2 dt^2 \left(1 - \frac{g^2 t^2}{c^2}\right)$$

where $g, c$ are constants. Determine the geodesics which correspond to it. Show that the family of geodesics which pass through $(0, 0, 0, 0)$ at $s = 0$ are $x = -\frac{1}{2}gt^2 + at$, $y = bt$, $z = kt$ where $a, b, k$ are constants. By transforming from $(x, y, z, t)$ to $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ where $\bar{x} = x + \frac{1}{2}gt^2$, $\bar{y} = y$, $\bar{z} = z$, $\bar{t} = t$ suggest a possible physical interpretation for $V_4$.

7.9    A certain $V_4$ has the line-element

$$ds^2 = u \, dt^2 - \frac{1}{u} dx^2 - dy^2 - dz^2$$

where $u = 1 - 2gx$ and $g$ is a positive constant. Show that a family of geodesics which pass through $(0,0,0,0)$ at $s = 0$ is given by

$$u^{1/2} \cosh gt = 1, \qquad y = z = 0.$$

By considering the line-element under the transformation

$$\bar{x} = \frac{u^{1/2} \cosh gt - 1}{g}$$

$$\bar{y} = y, \qquad \bar{z} = z,$$

$$\bar{t} = \frac{u^{1/2} \sinh gt}{g}$$

discuss this geodesic in the $(\bar{x}, \bar{y}, \bar{z}, \bar{t})$ coordinate system.

7.10   Verify that

$$1 + \frac{gx}{c^2} = \operatorname{sech} \frac{gt}{c}, \qquad y = 0, \qquad z = 0,$$

are the equations of a geodesic in a $V_4$ for which

$$ds^2 = \left(1 + \frac{gx}{c^2}\right)^2 c^2 dt^2 - ((dx)^2 + (dy)^2 + (dz)^2),$$

$g$ and $c$ being constants. In this $V_4$ particles follow geodesics. A particle is at rest at $x = y = z = 0$ at time $t = 0$. Show that in the subsequent motion $dx/dt$ attains a maximum of $c/2$.

7.11   A particular $V_4$ has the line-element

$$ds^2 = \exp\left(\frac{2t}{a}\right)(dx^2 + dy^2 + dz^2) - dt^2,$$

where $a$ is a positive constant. A null geodesic passes through $x = x_0 > 0$, $y = 0, z = 0$ at $t = t_0$ in the direction of the negative $x$-axis. Prove that if

$$x_0 > a \exp\left(\frac{-t_0}{a}\right)$$

the null geodesic will never pass through the origin.

7.12   If $V_4$ has the line-element

$$ds^2 = \frac{1}{r^2} dt^2 - \frac{1}{r^2} dr^2 - d\theta^2 - \sin^2 \theta d\phi^2$$

show that those geodesics for which $r = r_0$, $dr/ds = 0$, and $\dfrac{d\theta}{ds} = 0$ at $s = 0$ always move away from the origin.

7.13   Determine the geodesics corresponding to

$$ds^2 = z \, dt^2 - \frac{1}{z} dz^2 - dx^2 - dy^2.$$

7.14   Explain how the identity (2.17) can be used in order to compute the Christoffel symbols of the second kind for a given metric.

7.15   If $\phi$ is a scalar field in $V_n$ and

$$H_{ij} = \phi_{|i|j} - g_{ij}g^{rs}\phi_{|r|s}$$

prove that

$$H_{ij|k}g^{jk} = R_i^j \phi_{|j}.$$

*7.16  Prove that

$$R^{ji}{}_{kh|i|l} + R^{ji}{}_{lk|i|h} + R^{ji}{}_{hl|i|k} = R^{ij}{}_{kh}R_{il} + R^{ij}{}_{lk}R_{ih} + R^{ij}{}_{hl}R_{ik}.$$

7.17 Let the metric tensor of $V_n$ be such that

$$R_l{}^j{}_{kh|j} = 0.$$

Show that under these circumstances
(a) $R_{mk|h} = R_{mh|k}$,
(b) $R = $ constant,
(c) $R^{mj}{}_{kh}R_{ml} + R^{mj}{}_{lk}R_{mh} + R^{mj}{}_{hl}R_{mk} = 0.$

7.18 In $V_4$ show that
(a) $\varepsilon^{ijkh}R_{khaj} = 0$,
(b) $\varepsilon^{ijkh}R_{ijab|k} = 0$,
(c) if $F_{ij} = -F_{ji}$ then

$$\varepsilon^{ijkh}F_{ij|k} = 0 \quad \text{if and only if} \quad F_{ij|k} + F_{ki|j} + F_{jk|i} = 0.$$

7.19 (a) If $\psi$ is a scalar field of $V_n$ with positive definite metric, show that

$$g^{ij}\psi_{|i|j} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i}\left(\sqrt{g}\, g^{ij} \frac{\partial \psi}{\partial x^j}\right)$$

where $g = \det(g_{ij})$.
(b) If $ds^2 = dx^2 + dy^2 + dz^2$ show that $g^{ij}\psi_{|i|j} = 0$ reduces to

$$\nabla^2\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} = 0.$$

(c) By considering the transformation $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ show that the line-element in (b) becomes

$$ds^2 = dr^2 + r^2\,d\theta^2 + r^2\sin^2\theta\,d\phi^2,$$

and hence evaluate $\nabla^2\psi$ in spherical polars.

7.20 If $R_{ihjk} + R_{ijhk} = 0$ show that $R_{ijkl} = 0$.

7.21 Prove that

$$R_{ijkl}R^{ijkl} = 2R_{ijkl}R^{ikjl}$$

where

$$R^{ijkl} = g^{ia}g^{jb}g^{kc}g^{ld}R_{abcd},$$

and that

$$R^{ij}{}_{kh|i|j} = 0.$$

7.22 In $V_n$ a vector $B_i$ satisfies $g^{ij}B_iB_j = 0$ together with $B_{i|j} + B_{j|i} = 0$. If $B^i = g^{ij}B_j$ show that
(a) $B^i{}_{|i} = 0$,
(b) $B_{i|j}B^i = 0$,
(c) $B_{i|j}B^j = 0$,
(d) if $F_{ij} = B_{i|j} - B_{j|i}$ then $F_{ij}B^i = 0$,
(e) $g^{ih}F_{ij|h} = 2R^k_jB_k$.

*7.23 If

$$B^{hk} = g^{hj}g^{ki}[2R_{jlim}R^{ml} + g^{ml}R_{ij|l|m} - R_{|i|j}] - \tfrac{1}{2}g^{hk}(R^i_mR^m_i - g^{lm}R_{|l|m}),$$

show that
(a) $B^{hk} = B^{kh}$,
(b) $B^{hk}{}_{|k} = 0$.

7.24 In $V_3$, $A^{ij}$ is defined by

$$A^{ij} = \varepsilon^{ihk}(R^j_h - \tfrac{1}{4}\delta^j_h R)_{|k}.$$

Show that

$$A^{ij} = A^{ji}, \quad A^{ij}{}_{|j} = 0, \quad g_{ij}A^{ij} = 0.$$

7.25 If $V_2$ is such that $R = 0$ prove that $V_2$ is flat.
If $V_3$ is such that $R_{ij} = 0$ prove that $V_3$ is flat.

7.26 Obtain all the Christoffel symbols of the second kind for

$$ds^2 = [(du)^2 + (dv)^2]f(u + v)$$

where $f$ is a function of $u + v$ alone.
Hence solve

$$R_{ijkl} = 0$$

for $f$.

7.27 If $ds^2 = (dx^1)^2 + (x^1)^2(dx^2)^2 + (x^1)^2\sin^2 x^2(dx^3)^2$ show that

$$R_{ijkl} = 0.$$

7.28 If, in $V_2$, $ds^2 = (dx^1)^2 + H^2(dx^2)^2$, where $H = H(x^1, x^2)$, show that

$$R_{1212} = -H\frac{\partial^2 H}{(\partial x^1)^2}.$$

Under what conditions is $R_{ijhk} = 0$?

7.29 If $V_2$ is flat and

$$ds^2 = f(r)[(dx^1)^2 + (dx^2)^2]$$

where $r = [(x^1)^2 + (x^2)^2]^{1/2}$, find $f$.

7.30 A particular $V_4$ has line-element

$$ds^2 = e^\lambda(dx^1)^2 + (dx^2)^2 + (dx^3)^2 + e^\mu(dx^4)^2$$

where $\lambda, \mu$ are functions of $x^1$ alone. Prove that $V_4$ is flat if and only if

$$2\mu'' - \lambda'\mu' + (\mu')^2 = 0.$$

Obtain $\mu$ if (a) $\lambda = 0$, (b) $\lambda = -\mu$. (See Problem 5.26.)

7.31 A tensor $C_{j\ hm}^{\ l}$ is defined, for $n > 2$, by

$$C_{j\ hm}^{\ l} = R_{j\ hm}^{\ l} + \frac{1}{n-2}(R_h^l g_{jm} + R_{jm}\delta_h^l - R_{hj}\delta_m^l - R_m^l g_{hj})$$

$$+ \frac{R}{(n-1)(n-2)}(g_{hj}\delta_m^l - g_{jm}\delta_h^l).$$

Show that, if

$$C_{jkhm} = g_{kl}C_{j\ hm}^{\ l} \quad \text{then} \quad C_{jlhm} = -C_{ljhm} = -C_{jlmh},$$

$$C_{hjkl} + C_{ljhk} + C_{kjlh} = 0 \quad \text{and} \quad C_{i\ kj}^{\ j} = 0.$$

($C_{j\ hm}^{\ l}$ are called the components of the Weyl conformal curvature tensor; Weyl [1].)

7.32 Prove that if $n = 3$ then $C_{i\ kl}^{\ j} = 0$.

7.33 Show that if $L_{ij} = [g_{ij}R - 2(n-1)R_{ij}]/2(n-1)(n-2)$ then

$$C_{h\ il}^{\ k} = R_{h\ il}^{\ k} + \delta_l^k L_{hi} - \delta_i^k L_{hl} + g^{jk}g_{ih}L_{jl} - g^{jk}g_{lh}L_{ji}$$

and

$$C_{h\ il|j}^{\ j} = (n-3)(L_{hl|i} - L_{hi|l}).$$

*7.34 (a) Let the points of two Riemannian spaces $V_n$, $\tilde{V}_n$ be in one-to-one correspondence, corresponding points having the same coordinates $x^i$, the respective metrics being $g_{ij}(x^h)$ and $\tilde{g}_{ij}(x^h)$. If there exists a scalar $\sigma = \sigma(x^h)$ for which $\tilde{g}_{ij} = e^{2\sigma}g_{ij}$, $V_n$ and $\tilde{V}_n$ are said to be *conformal*. Show that if $V_n$ and $\tilde{V}_n$ are conformal then

$$\tilde{\gamma}_{ij}^h = \gamma_{ij}^h + \delta_i^h \sigma_{|j} + \delta_j^h \sigma_{|i} - g_{ij}g^{hk}\sigma_{|k},$$

where $\tilde{\gamma}_{ij}^h$ are the Christoffel symbols of the second kind formed with respect to $\tilde{g}_{ij}$. If $\tilde{L}_{ij}$ and $\tilde{C}_{i\ kl}^{\ j}$ are constructed from $\tilde{g}_{ij}$ in the same way as $L_{ij}$ and $C_{i\ kl}^{\ j}$ are constructed from $g_{ij}$ (see Problem 7.33) show that

$$\tilde{C}_{i\ kl}^{\ j} = C_{i\ kl}^{\ j} \quad \text{and} \quad L_{ih|j} - L_{ij|h} + C_{i\ jh}^{\ l}\sigma_{|l} = \tilde{L}_{ih;j} - \tilde{L}_{ij;h}$$

where the semicolon denotes covariant differentiation with respect to $\tilde{g}_{ij}$.

(b) If $V_n$ and $\tilde{V}_n$ are conformal and $\tilde{V}_n$ is flat then $V_n$ is called conformally flat. Show that if $V_n$ is conformally flat then

$$C_{i\ kl}^{\ j} = 0 \quad \text{and} \quad L_{ih|j} - L_{ij|h} = 0.$$

Discuss the dependence of these two conditions, paying particular attention to the cases $n = 3$ and $n > 3$ (see Problems 7.32 and 7.33).

(c) Consider a $V_n$ for which

$$C_{i\ kl}^{\ j} = 0 \quad \text{and} \quad L_{ih|j} - L_{ij|h} = 0.$$

Show that these conditions guarantee the existence of a scalar $\sigma$ for which

$$\sigma_{|h|i} - \sigma_{|i}\sigma_{|h} + \tfrac{1}{2}(g^{rs}\sigma_{|r}\sigma_{|s}) = -L_{ih}.$$

By defining

$$\tilde{g}_{ij} = e^{2\sigma}g_{ij}$$

show that $\tilde{V}_n$ is flat. Hence establish that $V_n$ is conformally flat if and only if

(a) $L_{ih|j} - L_{ij|h} = 0$   for $n = 3$,

(b) $C_{i\ kl}^{\ j} = 0$   for $n > 3$.

7.35 In the notation of Problem 7.34, if $F_{ij} = -F_{ji}$ is an arbitrary tensor field in $V_4$ and $\tilde{V}_4$ show that

$$(\sqrt{g}\ g^{ih}g^{jk}F_{hk})_{|i} = (\sqrt{\tilde{g}}\ \tilde{g}^{ih}\tilde{g}^{jk}\tilde{F}_{hk})_{;i}$$

where $g = |\det(g_{ij})|$ and $\tilde{g} = |\det(\tilde{g}_{ij})|$.

7.36 If, in $V_4$, $*R_{\ \ kl}^{ij} = \varepsilon^{ijab}R_{abkl}$ show that $*R_{\ \ kl}^{ij}$ has the properties (4.2.6).

7.37 Show that, if $n = 4$ (see Problem 4.9),

(a) $*R_{\ \ kl}^{ij}*R_{\ \ ih}^{kl} = \tfrac{1}{4}\delta_h^j(*R_{\ \ tu}^{rs}*R_{\ \ rs}^{tu})$;

(b) $C_{\ \ kl}^{ij}C_{\ \ ih}^{kl} = \tfrac{1}{4}\delta_h^j(C_{\ \ tu}^{rs}C_{\ \ rs}^{tu})$

where $C_{\ \ kl}^{ij} = g^{ih}C_{h\ kl}^{\ j}$ (Bach [1], Lanczos [1], Lovelock [1, 4]).

7.38 If $\phi^i$ is a vector field in $V_4$ for which $C_{i\ kl}^{\ j}\phi^i = 0$ show that either $g_{ij}\phi^i\phi^j = 0$ or $V_4$ is conformally flat (Brinkmann [1], Lovelock [4]). (See Problem 4.10.)

*7.39 The Bel-Robinson tensor $T^{ijkh}$ is defined, in $V_4$, by

$$T^{ijkh} = R_{\ \ \ b}^{iak}R_{\ \ \ a}^{hbj} + R_{\ \ \ b}^{iah}R_{\ \ \ a}^{kbj} - \tfrac{1}{2}g^{ij}R_{\ \ \ b}^{alh}R_{\ \ al}^{kb}$$

where

$$R_{\ \ al}^{kb} = g^{kc}R_{c\ al}^{\ b}, \quad R_{\ \ \ l}^{kbd} = g^{da}R_{\ \ al}^{kb}.$$

Show that

$$T^{ijkh} = T^{jikh} = T^{ijhk}.$$

Hence show that if $R_{ij} = 0$ then

(a) $T^{ijkh}$ is totally symmetric,

(b) $T^{ijkh}g_{ij} = 0$,

(c) $T^{ijkh}_{\ \ \ \ |i} = 0$

(Bel [1], Robinson [1]).

7.40 Prove that $G_j^i = -\tfrac{1}{4}\delta_{jrs}^{iab}R_{\ \ ab}^{rs}$.

Hence show that if $n = 4$

$$G_4^4 = -(R_{\ \ 12}^{12} + R_{\ \ 23}^{23} + R_{\ \ 31}^{31})$$

$$G_1^4 = R_{\ \ 21}^{24} + R_{\ \ 31}^{34},$$

and obtain the remaining components of $G_j^i$.

7.41 Consider a particular $V_3$ with line-element

$$ds^2 = (dx^1)^2 + (dx^2)^2 - (dx^3)^2$$

and the hypersurface characterized by

$$x^1 = u^1, \qquad x^2 = u^2, \qquad x^3 = \sqrt{(u^1)^2 + (u^2)^2}.$$

Show that all points on this surface are such that

$$\det(g_{\alpha\beta}) = 0.$$

7.42 Show that if $g_{ij}$ is positive definite then so is $g_{\alpha\beta}$, defined by (4.5).

7.43 In $E_4$ with coordinates $x^i$ ($i = 1, 2, 3, 4$) and line-element

$$ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2 + (dx^4)^2$$

a surface $S$ is defined by

$$\begin{aligned}
x^1 &= \cos u^1 \\
x^2 &= \sin u^1 \cos u^2 \\
x^3 &= \sin u^1 \sin u^2 \cos u^3 \\
x^4 &= \sin u^1 \sin u^2 \sin u^3
\end{aligned}$$

(hypersphere of radius 1). Compute $B_\alpha^i$, $g_{\alpha\beta}$, $N^i$, $\Omega_{\alpha\beta}$, and $R_{\alpha\beta\lambda\mu}$ for $S$.

7.44 In $E_3$ with line-element $ds^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$, the surface area $A$ of the hypersurface $x^i = x^i(u^\alpha)$ ($\alpha = 1, 2$) is defined by

$$A = \iint_R \sqrt{\det g_{\alpha\beta}} \, du^1 \, du^2$$

where $R$ is the appropriate region in the $u^1$, $u^2$ plane.

(a) If the surface is characterized by

$$x^1 = u^1, \qquad x^2 = u^2, \qquad x^3 = f(u^1, u^2),$$

show that

$$A = \iint_R \sqrt{1 + (\partial f/\partial u^1)^2 + (\partial f/\partial u^2)^2} \, du^1 \, du^2.$$

(b) Under transformations of the type $\tilde{u}^\alpha = \tilde{u}^\alpha(u^\beta)$ show that $A$ is a scalar. Interpret this result.

7.45 In $E_3$ show that a surface of revolution is characterized by

$$\begin{aligned}
x^1 &= f(u^1)\cos u^2 \\
x^2 &= f(u^1)\sin u^2 \\
x^3 &= u^1
\end{aligned}$$

($0 \le u^2 \le 2\pi$). Evaluate its surface area.

7.46 In $E_3$ a torus is defined by

$$\begin{aligned}
x^1 &= (b + a \sin u^1)\cos u^2 \\
x^2 &= (b + a \sin u^1)\sin u^2 \\
x^3 &= a \cos u^1,
\end{aligned}$$

where $a$, $b$ are constants ($a < b$), $0 \le u^1 \le 2\pi$, $0 \le u^2 \le 2\pi$. What is the surface area?

*7.47 Show that if $n > 4$ and $\det(\Omega_{\alpha\beta}) \neq 0$ then (5.28) imply (5.32) (Thomas [2], Berry [1]).

*7.48 (a) If

$$A_i^j(r) = \delta_{ii_1\cdots i_r}^{jj_1\cdots j_r} \delta_{\beta_1\cdots\beta_r}^{\alpha_1\cdots\alpha_r} B_{\alpha_1}^{i_1} \cdots B_{\alpha_r}^{i_r} B_{j_1}^{\beta_1} \cdots B_{j_r}^{\beta_r}$$

show that

$$A_i^j(r) = \frac{r!(n-1)!}{(n-r-1)!} \delta_i^j - r^2 B_\alpha^j B_{j_1}^\alpha A_i^{j_1}(r-1).$$

Hence show, by induction, that

$$A_i^j(r) = \frac{r!(n-1)!}{(n-r-1)!} \delta_i^j - \frac{rr!(n-2)!}{(n-r-1)!} B_\alpha^j B_i^\alpha.$$

(b) If

$$N_i = \lambda \varepsilon_{ii_1\cdots i_{n-1}} \varepsilon^{\alpha_1\cdots\alpha_{n-1}} B_{\alpha_1}^{i_1} \cdots B_{\alpha_{n-1}}^{i_{n-1}} \sqrt{\frac{|\det g_{ij}|}{|\det g_{\alpha\beta}|}}$$

show that $g^{ij} N_i N_j = 1$ if $\lambda = 1/(n-1)!$. Hence, with this choice of $\lambda$ and by considering $N_i g^{jk} N_j$ show that

$$B_\alpha^j B_i^\alpha = \delta_i^j - g^{jk} N_i N_k.$$

7.49 Prove that in $E_3$ on the surface of a cylinder the geodesics are helices.

*7.50 Prove that the geodesics on the surface of a sphere in $E_3$ are great circles, that is, they lie in a plane through the center of the sphere. *Hint:* show that, in $\theta$, $\phi$ coordinates, one of the geodesic equations integrates to

$$\sin^2 \theta \, \frac{d\phi}{ds} = \sin \alpha$$

where $\alpha$ is a constant, and this gives rise to

$$\frac{d\phi}{d\theta} = \frac{\sin \alpha}{\sin \theta (\sin^2 \theta - \sin^2 \alpha)^{1/2}}.$$

In this differential equation let $u = \tan \alpha \cot \theta$ to find

$$\cos(\phi - \beta) = \frac{\tan \alpha}{\tan \theta},$$

and write the latter in terms of $x, y, z$.

*7.51 (Mercator's projection)

(a) Find the geodesics corresponding to

$$ds^2 = \operatorname{sech}^2 y (dx^2 + dy^2).$$

(b) In $E_3$ a hemisphere of unit radius (coordinates $\theta$, $\phi$) is mapped onto the $x, y$ plane by

$$\begin{aligned}
x &= \phi \\
y &= \ln(\cot \tfrac{1}{2}\theta).
\end{aligned}$$

Calculate the metric for the hemispherical surface in terms of $x$ and $y$ and show that the great circles are represented by the curves

$$\sinh y = \alpha \sin(x + \beta),$$

$\alpha$, $\beta$ constants.

7.52 The coefficients of the $r$th fundamental form $C_{(r)\alpha\beta}$ of a hypersurface $V_{n-1}$ of a $V_n$ are defined by the recursion formula

$$C_{(r)\alpha\beta} = g^{\varepsilon\gamma}\Omega_{\alpha\varepsilon}C_{(r-1)\gamma\beta} \qquad (r = 2, \ldots, n),$$

where $C_{(1)\alpha\beta} = g_{\alpha\beta}$, $C_{(2)\alpha\beta} = \Omega_{\alpha\beta}$. With each of these one may associate the scalar $M_{(r)}$ defined by

$$M_{(r)} = g^{\alpha\beta}C_{(r+1)\alpha\beta} \qquad (r = 0, \ldots, n-1).$$

Show that the fundamental invariants of $V_{n-1}$ as defined by (5.18) can be expressed explicitly in the form

$$2! H_{(2)} = M_{(1)}^2 - M_{(2)},$$

$$3! H_{(3)} = M_{(1)}^3 - 3M_{(1)}M_{(2)} + 2! M_{(3)},$$

$$4! H_{(4)} = M_{(1)}^4 - 6M_{(1)}^2 M_{(2)} + 3M_{(2)}^2 + 8M_{(1)}M_{(3)} - 3! M_{(4)},$$

and so forth. By means of the equation (5.28) and the above expression for $H_{(2)}$ show that

$$2H_{(2)} = \bar{R} - R + 2R_{lh}N^l N^h,$$

where $\bar{R}$, $R$ denote the scalar curvatures of $V_{n-1}$ and $V_n$, respectively. Show that if the embedding space is Euclidean,

$$4! H_4 = R^{\alpha\beta\gamma\varepsilon}R_{\alpha\beta\gamma\varepsilon} + \bar{R}^2 - 4R^{\alpha\gamma}R_{\alpha\gamma}$$

(Rund [5]).

7.53 By means of Newton's formula show that, for a hypersurface $V_{n-1}$ of a $V_n$, the quantities $M_{(r)}$ as defined in Problem 7.52 are related to the invariants (5.18) according to

$$(-1)^{r+1}rH_{(r)} = \sum_{s=1}^{r-1}(-1)^s H_{(s)}M_{(r-s)},$$

in which $H_{(0)} \equiv 1$. Using the Cayley-Hamilton theorem show that the coefficient of the $n$th fundamental form of $V_{n-1}$ is given by

$$C_{(n)\alpha\beta} + \sum_{s=1}^{n-1}(-1)^s H_{(s)}C_{(n-s)\alpha\beta} = 0.$$

(*Note*: this is a generalization of the well-known theorem in the classical differential geometry of surfaces according to which the third fundamental form may be expressed as a linear combination of the first and second fundamental forms.) (Rund [7]).

7.54 In the notation of Problem 7.52, show that

$$\frac{\partial H_{(s)}}{\partial u^\alpha} = \sum_{t=1}^{s}(-1)^{t+1}t^{-1}H_{(s-t)}\frac{\partial M_{(t)}}{\partial u^\alpha}.$$

7.55 The coefficients of the associated fundamental forms of a $V_{n-1}$ in a $V_n$ are defined by

$$(-1)^{r+1}P_{(r)\alpha\beta} = \sum_{s=0}^{r-1}(-1)^s H_{(s)}C_{(r-s)\alpha\beta} \qquad (r = 1, \ldots, n),$$

in the notation of Problem 7.52.

Show that, if $V_n$ is a space of constant curvature,

$$g^{\alpha\lambda}P_{(r)\lambda\beta\|\alpha} = 0$$

(Rund [6]).

7.56 Show that the $r$th curvature of $V_m$ is given by

$$H_{(r)} = \delta^{\alpha_1\cdots\alpha_r}_{\beta_1\cdots\beta_r}\Omega^{\beta_1}_{\alpha_1}\cdots\Omega^{\beta_r}_{\alpha_r}/r!$$

Hence show that the curvatures $H_{(2)}, H_{(4)}, \ldots, H_{(2p)}$ of a hypersurface embedded in $E_n$, with $2p + 1 \leq n$, are intrinsic quantities of the hypersurface (Lovelock [6]).

7.57 If $X^i$, $Y^i$, $Z^i$, $W^i$ are components of arbitrary type $(1, 0)$ tensors the notation $R(X, Y)Z$ is frequently used to denote $R_i{}^j{}_{kl}X^k Y^l Z^i$ while $\langle R(X, Y)Z, W\rangle$ represents $R_i{}^j{}_{kl}X^k Y^l Z^i g_{jh}W^h$ (Nelson [1]). Show that

$$R(X, Y)Z = -R(Y, X)Z,$$

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0,$$

$$\langle R(X, Y)Z, W\rangle = -\langle R(X, Y)W, Z\rangle,$$

$$\langle R(X, Y)Z, W\rangle = \langle R(Z, W)X, Y\rangle.$$

7.58 If, in $V_n$, $A_i$ satisfies $A_{i\|j} + A_{j\|i} = 0$ (the Killing equation) show that, along any geodesic, $A_i dx^i/ds$ is constant. Conversely prove that if $A_i dx^i/ds$ is constant along any geodesic then $A_i$ satisfies Killing's equation.

7.59 In a particular $V_4$, $A_i$ satisfies the Killing equation and $A_i A^i = $ constant ($\neq 0$), where $A^i = g^{ij}A_j$. If $R_{ij} = 0$ show that
(a) $A^i{}_{\|ji} = 0$,
(b) $A_{i\|j}A^i{}_{\|h}g^{hj} = 0$,
(c) $A_i B^i = 0$ where $B^i = \varepsilon^{ijkl}A_{j\|k}A_l$,
(d) $B^i B^j g_{ij} = 0$.
If $A_i$ has the additional property that all null vectors orthogonal to it vanish identically, prove that $R_i{}^j{}_{kl}A_j = 0$ and hence that $V_4$ is flat. (See Problem 7.38.)

7.60 If, in a particular $V_4$, there exists a non-null vector $A_i$ for which

$$C_{ijkl}A^i A^k = 0$$

and

$$\varepsilon^{abkl}C_{ijkl}A_a A^j = 0,$$

where $A^i = g^{ij}A_j$ and $C_{ijkl}$ is defined in Problem 7.31, show that $V_4$ is conformally flat.

# 8

# INVARIANT VARIATIONAL PRINCIPLES AND PHYSICAL FIELD THEORIES

Invariant variational principles have had a profound influence on physical field theories, and this final chapter is devoted to an introduction to the subject.

A physical field is described by the so-called field functions (e.g., the components of a vector field), and it is usually assumed that the field equations that govern the behavior of the field are identical with the Euler-Lagrange equations of a given problem in the calculus of variations. The action integral of the latter is supposed to be invariant under coordinate transformations; this implies that the corresponding Lagrangian is a scalar density. However, for any given type of field, this invariance requirement must be augmented by an additional assumption concerned with invariance properties, namely, one which specifies the transformation properties of the field functions. These two invariance requirements, taken in conjunction, severely restrict the classes of admissible Lagrangians and hence also the types of acceptable field equations.

In this chapter three such field theories† are considered, of which the first two are concerned with field functions which are tensor fields of type $(0, 1)$ and $(0, 2)$, respectively, while the third is a combination of the previous

† References to other field theories which have been considered can be found in a survey article by Rund and Lovelock [1].

cases. These examples have been selected because they contain Maxwell's equations, the Einstein field equations, and the Einstein-Maxwell equations, as special cases. The techniques developed in this way are then used to significantly strengthen an earlier result due to E. Cartan [1] and Weyl [2], which is frequently used to motivate the field equations of Einstein.

Although oriented toward the needs of relativists, this chapter requires no previous knowledge of relativity. Nevertheless, an acquaintance with the elements of general relativity as covered in, for example, Adler, Bazin, and Schiffer [1] would be beneficial.

## 8.1  INVARIANT FIELD THEORIES

In this section we discuss the role played by invariant variational principles in the theory of physical fields, with particular reference to the electromagnetic and gravitational fields. We therefore adopt a notation which is more appropriate to these applications than that used in Chapter 6. In general we shall be concerned with action integrals of the type

$$I = \int_G L(x^h, \rho^A(x^h), \rho^A_{,h}, \rho^A_{,hk}) \, d(x) \tag{1.1}$$

where the Lagrangian $L$ is a given function (assumed to be of class $C^4$ in all its arguments), $G$ is a prescribed simply connected region of our given manifold $X_n$ (referred to local coordinates $x^h$), and $\rho^A(x^h)$ $(A = 1, \ldots, M)$ denotes the dependent variables with

$$\rho^A_{,h} = \frac{\partial \rho^A}{\partial x^h}, \quad \rho^A_{,hk} = \frac{\partial^2 \rho^A}{\partial x^h \, \partial x^k}$$

and

$$d(x) = dx^1 \cdots dx^n.$$

A comparison of (1.1) with (6.7.1) shows that the role played by the set $(t^\alpha, x^h(t^\alpha))$ in Chapter 6 is here played by the set $(x^h, \rho^A(x^h))$. In this notation the Euler-Lagrange expression (6.7.20) is

$$E_A(L) = \frac{d}{dx^h} \left[ \frac{\partial L}{\partial \rho^A_{,h}} - \frac{d}{dx^k} \left( \frac{\partial L}{\partial \rho^A_{,hk}} \right) \right] - \frac{\partial L}{\partial \rho^A}, \tag{1.2}$$

and the Euler-Lagrange equations (6.7.19) are

$$E_A(L) = 0. \tag{1.3}$$

It is usually assumed that the integral (1.1) is invariant under (proper) coordinate transformations

$$\bar{x}^h = \bar{x}^h(x^j), \tag{1.4}$$

which implies that $L$ must be a scalar density, that is, that the transform of $L$ under (1.4) is given by

$$\bar{L}(\bar{x}^h, \bar{\rho}^A, \bar{\rho}^A_{,h}, \bar{\rho}^A_{,hk}) = JL(x^h, \rho^A, \rho^A_{,h}, \rho^A_{,hk}), \tag{1.5}$$

where

$$J = \det(J^h_k) \quad (> 0), \tag{1.6}$$

and

$$J^h_k = \frac{\partial x^h}{\partial \bar{x}^k}. \tag{1.7}$$

We henceforth restrict our considerations to such Lagrangians.

Up to this stage no assumptions have been made concerning the transformation properties of $\rho^A$ under (1.4), and in order to proceed further we are forced to discuss specific cases. We shall primarily be concerned with two such cases: in the first of these $\rho^A$ will represent the components of a type (0, 1) tensor field (a covariant vector field), while in the second case they will represent the components of a type (0, 2) symmetric tensor field (a metric field). These particular examples have been selected, first because they serve to illustrate the typical techniques used to analyze the problem under consideration, and second because of their importance when applied to certain physical field theories.

## 8.2    VECTOR FIELD THEORY

In this first case we consider scalar densities of the type

$$L = L(x^h, \psi_h, \psi_{h,k}, g_{hk}) \tag{2.1}$$

where $\psi_h$ are the components of a covariant vector field (and here play the role of $\rho^A$) and $g_{hk}$ are the components of a symmetric type (0, 2) tensor field for which $g = |\det(g_{hk})| \neq 0$. It is assumed of the $g_{hk}$ that they are arbitrary but preassigned whereas the $\psi_h$ are the functions to be determined by the variational method (Rund [3]).

By considering the particular transformation

$$\bar{x}^h = x^h + a^h \tag{2.2}$$

where $a^h$ are constants, it is easily seen that (1.5) and (2.1) imply

$$L = L(\psi_h, \psi_{h,k}, g_{hk}). \tag{2.3}$$

For convenience we introduce the notation

$$\left. \begin{aligned} \Psi^i &= \frac{\partial L}{\partial \psi_i}, \\ \Psi^{ij} &= \frac{\partial L}{\partial \psi_{i,j}}, \\ \Lambda^{ij} &= \frac{\partial L}{\partial g_{ij}}, \end{aligned} \right\} \tag{2.4}$$

where, in the latter, we have regarded the arguments $g_{ij}$ in (2.3) as representing $\frac{1}{2}(g_{ij} + g_{ji})$ in which $g_{ij}$ and $g_{ji}$ are to be considered as independent. As a consequence of this $\Lambda^{ij}$ is symmetric. Our first task will be to discuss the tensorial character of $\Psi^i$, $\Psi^{ij}$, and $\Lambda^{ij}$.

Under the transformation (1.4) the arguments of (2.3) transform as follows:

$$\left. \begin{aligned} \bar{\psi}_h &= J^i_h \psi_i, \\ \bar{\psi}_{h,k} &= J^i_{h\,k} \psi_i + J^i_h J^j_k \psi_{i,j}, \\ \bar{g}_{hk} &= J^i_h J^j_k g_{ij}, \end{aligned} \right\} \tag{2.5}$$

where $J^h_k$ is defined by (1.7) and

$$J^i_{h\,k} = \frac{\partial J^i_h}{\partial \bar{x}^k} = \frac{\partial^2 x^i}{\partial \bar{x}^h\, \partial \bar{x}^k}. \tag{2.6}$$

If we take (2.3) and (2.5) into account then in this case (1.5) reduces to the following identity in $\psi_i$, $\psi_{i,j}$, $g_{ij}$, $J^i_h$, and $J^i_{h\,k}$:

$$\bar{L}(J^i_h \psi_i, J^i_{h\,k} \psi_i + J^i_h J^j_k \psi_{i,j}, J^i_h J^j_k g_{ij}) = JL(\psi_i, \psi_{i,j}, g_{ij}). \tag{2.7}$$

Differentiating (2.7) with respect to $\psi_i$, $\psi_{i,j}$, and $g_{ij}$ we obtain, by virtue of (2.4),

$$\bar{\Psi}^{hk} J^i_{h\,k} + \bar{\Psi}^h J^i_h = J\Psi^i, \tag{2.8}$$

$$\bar{\Psi}^{hk} J^i_h J^j_k = J\Psi^{ij}, \tag{2.9}$$

and

$$\bar{\Lambda}^{hk} J^i_h J^j_k = J\Lambda^{ij}, \tag{2.10}$$

respectively. The relations (2.9) and (2.10) clearly indicate that $\Psi^{ij}$ and $\Lambda^{ij}$ are the components of tensor densities of type (2, 0). However, in view of the term $\bar{\Psi}^{hk} J^i_{h\,k}$ in (2.8), it appears that $\Psi^i$ have no tensorial character. We shall return to this point shortly.

In order to exploit (2.7) further we shall eventually differentiate it with respect to $J^i_h$ and $J^i_{h\,k}$, which means that we will require the corresponding

derivatives of $\bar{\psi}_h$, $\bar{\psi}_{h,k}$, $\bar{g}_{hk}$, and $J$. Taking into account the appropriate symmetry relations we find, from (2.5),

$$\left.\begin{aligned}
\frac{\partial \bar{\psi}_h}{\partial J_t^s} &= \delta_h^t \psi_s, \\[4pt]
\frac{\partial \bar{\psi}_{h,k}}{\partial J_t^s} &= \psi_{i,s} J_h^i \delta_k^t + \psi_{s,i} J_k^i \delta_h^t, \\[4pt]
\frac{\partial \bar{g}_{hk}}{\partial J_t^s} &= g_{is}(J_k^i \delta_h^t + J_h^i \delta_k^t), \\[4pt]
\frac{\partial \bar{\psi}_{h,k}}{\partial J_{s\;t}^{\;r}} &= \frac{1}{2}\psi_r(\delta_h^s \delta_k^t + \delta_h^t \delta_k^s),
\end{aligned}\right\} \tag{2.11}$$

the derivatives of $\bar{\psi}_h$ and $\bar{g}_{hk}$ with respect to $J_{s\;t}^{\;r}$ vanishing identically. Furthermore, if $A_h^i$ denotes the inverse of $J_k^h$, that is,

$$A_h^i J_k^h = \delta_k^i, \tag{2.12}$$

then, by (1.3.5),

$$\frac{\partial J}{\partial J_t^s} = J A_s^t. \tag{2.13}$$

We first differentiate (2.7) with respect to $J_{s\;t}^{\;r}$, which, by (2.4) and (2.11), yields

$$\overline{\Psi}^{hk}\tfrac{1}{2}\psi_r(\delta_h^s \delta_k^t + \delta_h^t \delta_k^s) = 0.$$

In general $\psi_r \neq 0$, in which case the latter equation implies

$$\overline{\Psi}^{st} + \overline{\Psi}^{ts} = 0, \tag{2.14}$$

or, in view of the tensorial character of $\Psi^{st}$,

$$\Psi^{st} + \Psi^{ts} = 0. \tag{2.15}$$

This is the first set of conditions which $L$ must satisfy as a result of (2.7). A remarkable feature of this result is the fact that gradient fields, that is, fields for which $\psi_{i,j} = \psi_{j,i}$, must be excluded from our considerations. For the latter condition combined with (2.15) yields $\Psi^{ij} = 0$, in which case $L$ would depend solely on $\psi_j$ and $g_{jh}$.

If (2.14) is applied to (2.8), we see that the first term on the left-hand side of the latter vanishes identically, which implies that the $\Psi^h$ are the components of a tensor density of type $(1, 0)$.

We now differentiate (2.7) with respect to $J_t^s$, and use (2.4), (2.11), and (2.13) to obtain

$$\overline{\Psi}^h \delta_h^t \psi_s + \overline{\Psi}^{hk}(\psi_{i,s} J_h^i \delta_k^t + \psi_{s,i} J_k^i \delta_h^t) + \overline{\Lambda}^{hk} g_{is}(J_k^i \delta_h^t + J_h^i \delta_k^t) = JLA_s^t. \tag{2.16}$$

To the second term on the left-hand side we apply (2.14). This suggests the following notation:

$$F_{js} = \psi_{s,j} - \psi_{j,s}, \tag{2.17}$$

these quantities being the components of a skew-symmetric type $(0, 2)$ tensor. The relation (2.16) can thus be written in the form

$$\overline{\Psi}^t \psi_s + J_k^j \overline{\Psi}^{tk} F_{js} + 2\overline{\Lambda}^{tk} g_{js} J_k^j = JLA_s^t. \tag{2.18}$$

Equation (2.18) is an identity in $J_k^j$. Consequently, if we consider the particular transformation

$$\bar{x}^j = x^j, \tag{2.19}$$

in which case $J_h^j = \delta_h^j$, $A_s^t = \delta_s^t$, $J = 1$, the relation (2.18) reduces to

$$\Psi^t \psi_s + \Psi^{tj} F_{js} + 2\Lambda^{tj} g_{js} = \delta_s^t L, \tag{2.20}$$

which is the second (tensorial) condition to be satisfied by $L$. Equations (2.15) and (2.20) are usually called *invariance identities*.

Now that we have exhausted the information which we can obtain from (2.7) we turn to the Euler-Lagrange expression (1.2) which, in this case, is

$$E^i(L) = \Psi^{ij}_{\ ,j} - \Psi^i. \tag{2.21}$$

By virtue of (2.5) and (2.9) this can be expressed in the form

$$E^i(L) = \Psi^{ij}_{\ |j} - \Psi^i, \tag{2.22}$$

where the vertical bar denotes covariant differentiation relative to the Christoffel symbols defined by the tensor $g_{ij}$. The relation (2.22) clearly indicates that the $E^i(L)$ are the components of a tensor density of type $(1, 0)$, so that the Euler-Lagrange equations (1.3), namely,

$$E^i(L) = 0, \tag{2.23}$$

are tensorial conditions.

From (2.21), (2.15), and (4.1.27) we obtain the identity

$$E^i(L)_{|i} = E^i(L)_{,i} = -\Psi^i_{\ ,i} = -\Psi^i_{\ |i}. \tag{2.24}$$

Consequently, if the Euler-Lagrange equations (2.23) are satisfied, then

$$\frac{\partial \Psi^i}{\partial x^i} = 0, \tag{2.25}$$

which is called the generalized Lorentz condition.

We now wish to obtain an additional identity which, upon (2.23) being satisfied, reduces to a divergence condition. If we differentiate (2.20) covariantly with respect to $x^t$, we find

$$2\Lambda^{tj}_{\ |t} g_{js} = L_{|s} - \Psi^t_{\ |t} \psi_s - \Psi^t \psi_{s|t} - \Psi^{tj}_{\ |t} F_{js} - \Psi^{tj} F_{js|t}. \tag{2.26}$$

Because of (2.17), (2.22), and (2.24), this can be expressed in the form

$$2\Lambda^{tj}_{\ |t} g_{js} = L_{|s} + E^t(L)_{|t}\psi_s + E^j(L)F_{js} - \Psi^t\psi_{t|s} - \Psi^{tj}F_{js|t}. \qquad (2.27)$$

By virtue of (2.15) and the identity

$$F_{js|t} + F_{tj|s} + F_{st|j} = 0 \qquad (2.28)$$

[which is an immediate consequence of (2.17)], we can express the last term on the right-hand side of (2.27) as follows:

$$-\Psi^{tj}F_{js|t} = -\tfrac{1}{2}\Psi^{tj}(F_{js|t} - F_{ts|j}) = \tfrac{1}{2}\Psi^{tj}F_{tj|s} = -\Psi^{tj}\psi_{t|js}.$$

Consequently (2.27) can be written in the form

$$2\Lambda^{tj}_{\ |t} g_{js} = E^t(L)_{|t}\psi_s + E^j(L)F_{js} + [L_{|s} - \Psi^t\psi_{t|s} - \Psi^{tj}\psi_{t|js}].$$

It is easily seen that the quantity in square brackets vanishes identically, so that we have

$$2\Lambda^{tr}_{\ |t} = g^{rs}\psi_s E^t(L)_{|t} + g^{rs}F_{js}E^j(L). \qquad (2.29)$$

From (2.29) we see that if (2.23) is satisfied, then

$$\Lambda^{tr}_{\ |t} = 0. \qquad (2.30)$$

The quantities $\Lambda^{tr}$ are usually called the components of the *energy-momentum tensor density* and, by (2.20), they can be expressed in the form

$$\Lambda^{tr} = \tfrac{1}{2}(g^{tr}L - \Psi^t g^{rs}\psi_s - g^{rs}\Psi^{tj}F_{js}). \qquad (2.31)$$

We conclude this section by considering the following particular choice for the Lagrangian:

$$L = \sqrt{g}\, g^{ih}(\tfrac{1}{2}g^{jk}F_{ij}F_{hk} + \mu\psi_i\psi_h), \qquad (2.32)$$

where $\mu$ is a constant. It is clearly a scalar density of the form (2.3) so that the analysis of this section is applicable. From (2.32) it is easily seen that

$$\Psi^{kj} = 2\sqrt{g}\, g^{ij}g^{hk}F_{ih} = 2\sqrt{g}\, F^{jk},$$

and

$$\Psi^k = 2\mu\sqrt{g}\, g^{jk}\psi_j.$$

Consequently, for this example, the identities (2.22) and (2.31) become

$$E^k(L) = 2(\sqrt{g}\, F^{jk})_{|j} - 2\mu\sqrt{g}\, g^{kj}\psi_j \qquad (2.33)$$

and

$$\Lambda^{tr} = -\sqrt{g}\,[F^{ti}F_{hi}g^{hr} - \tfrac{1}{4}g^{tr}(F_{ij}F^{ij})] - \mu\sqrt{g}\,[g^{ti}g^{rj}\psi_i\psi_j - \tfrac{1}{2}g^{tr}(g^{ij}\psi_i\psi_j)], \qquad (2.34)$$

respectively, while the condition (2.25) is

$$\mu\frac{\partial}{\partial x^k}(\sqrt{g}\, g^{kj}\psi_j) = 0. \qquad (2.35)$$

Consider the particular case when $g_{ij}$ represents the metric of Minkowski space-time, for which, by definition $g_{11} = g_{22} = g_{33} = -1$, $g_{44} = 1$, $g_{ij} = 0 \ (i \neq j)$. If $\mu = 0$, then (2.33) and (2.28) are Maxwell's equations while (2.34) is the well-known energy-momentum tensor of the electromagnetic field provided that we identify $\psi_i$ with the electromagnetic vector potential. If $\mu \neq 0$, then (2.35) reduces to the usual Lorentz condition.

## 8.3   METRIC FIELD THEORY

We now turn our attention to scalar densities of the form

$$L = L(x^h, g_{hk}, g_{hk,j}, g_{hk,jl}), \qquad (3.1)$$

where $g_{hk}$ are the components of a symmetric type (0, 2) tensor field with $g = |\det(g_{hk})| \neq 0$. Lagrangians of the kind (3.1) are considerably more complicated than those of kind (2.1) for two reasons: first, the field considered in (3.1), namely $g_{hk}$, satisfies a transformation law [under (1.4)] which is more complicated than that of the vector field, and second, the function (3.1) depends on second-order derivatives of the field whereas the Lagrangian (2.1) is dependent on first derivatives of the vector field. It will be seen that, as a result of these properties, the resulting calculations are substantially longer than those of the corresponding vector field case (Rund [3]).

By considering the particular transformation (2.2) it is again possible to show that (1.5) and (3.1) imply that

$$L = L(g_{hk}, g_{hk,j}, g_{hk,jl}), \qquad (3.2)$$

so that (1.5) can be written in the form

$$\bar{L}(\bar{g}_{hk}, \bar{g}_{hk,j}, \bar{g}_{hk,jl}) = JL(g_{hk}, g_{hk,j}, g_{hk,jl}). \qquad (3.3)$$

A fundamental role will again be played by the derivatives of $L$ with respect to its arguments, which we denote by

$$\Lambda^{ij} = \frac{\partial L}{\partial g_{ij}}, \qquad \Lambda^{ij,k} = \frac{\partial L}{\partial g_{ij,k}}, \qquad \Lambda^{ij,kh} = \frac{\partial L}{\partial g_{ij,kh}}, \qquad (3.4)$$

where the remark following (2.4), and its obvious generalization to $g_{ij,k}$ and $g_{ij,kh}$, is again applicable. Thus the following are immediate properties:

$$\Lambda^{ij} = \Lambda^{ji}, \qquad (3.5)$$

$$\Lambda^{ij,k} = \Lambda^{ji,k}, \qquad (3.6)$$

$$\Lambda^{ij,kh} = \Lambda^{ji,kh} = \Lambda^{ij,hk}. \tag{3.7}$$

As in the preceding section we first discuss the tensorial character of $\Lambda^{ij}$, $\Lambda^{ij,k}$, and $\Lambda^{ij,kh}$.

Under (1.4) the arguments of (3.2) transform as follows:

$$\bar{g}_{ij} = J_i^a J_j^b g_{ab}, \tag{3.8}$$

$$\bar{g}_{ij,k} = (J_{i\ k}^{\ a} J_j^b + J_i^a J_{j\ k}^{\ b}) g_{ab} + J_i^a J_j^b J_k^c g_{ab,c}, \tag{3.9}$$

$$\begin{aligned}
\bar{g}_{ij,kh} &= (J_{i\ kh}^{\ a} J_j^b + J_{i\ k}^{\ a} J_{j\ h}^{\ b} + J_{i\ h}^{\ a} J_{j\ k}^{\ b} + J_i^a J_{j\ kh}^{\ b}) g_{ab} \\
&+ (J_{i\ k}^{\ a} J_j^b J_h^c + J_i^a J_{j\ k}^{\ b} J_h^c + J_{i\ h}^{\ a} J_j^b J_k^c + J_i^a J_{j\ h}^{\ b} J_k^c + J_i^a J_j^b J_{k\ h}^{\ c}) g_{ab,c} \\
&+ J_i^a J_j^b J_k^c J_h^d g_{ab,cd}, \tag{3.10}
\end{aligned}$$

where

$$J_{i\ kh}^{\ a} = \frac{\partial J_{i\ k}^{\ a}}{\partial \bar{x}^h} = \frac{\partial^3 x^a}{\partial \bar{x}^i\, \partial \bar{x}^k\, \partial \bar{x}^h}.$$

If we differentiate (3.3) with respect to $g_{ab,cd}$, $g_{ab,c}$, and $g_{ab}$, taking into account (3.5)–(3.10), we find respectively that

$$J\Lambda^{ab,cd} = \overline{\Lambda}^{ij,kh} \frac{\partial \bar{g}_{ij,kh}}{\partial g_{ab,cd}} = \overline{\Lambda}^{ij,kh} J_i^a J_j^b J_k^c J_h^d, \tag{3.11}$$

$$J\Lambda^{ab,c} = \overline{\Lambda}^{ij,kh} \frac{\partial \bar{g}_{ij,kh}}{\partial g_{ab,c}} + \overline{\Lambda}^{ij,k} \frac{\partial \bar{g}_{ij,k}}{\partial g_{ab,c}}, \tag{3.12}$$

and

$$J\Lambda^{ab} = \overline{\Lambda}^{ij,kh} \frac{\partial \bar{g}_{ij,kh}}{\partial g_{ab}} + \overline{\Lambda}^{ij,k} \frac{\partial \bar{g}_{ij,k}}{\partial g_{ab}} + \overline{\Lambda}^{ij} \frac{\partial \bar{g}_{ij}}{\partial g_{ab}}. \tag{3.13}$$

The relation (3.11) establishes that the $\Lambda^{ab,cd}$ form the components of a type (4, 0) tensor density. However, because of the presence of the first term on the right-hand side of (3.12) and the first two terms on the right-hand side of (3.13), the quantities $\Lambda^{ab,c}$ and $\Lambda^{ab}$ have no tensorial character. Nevertheless, it is possible to introduce quantities closely related to these which are tensorial in character. We do this in the following indirect manner. Let $h_{ij}$ be the components of an arbitrary symmetric type (0, 2) tensor field, so that it and its derivatives transform in the manner indicated by (3.8)–(3.10) in which $g_{ij}$ is replaced by $h_{ij}$ throughout.

We now define a new quantity $F$ by †

$$F = \Lambda^{ab,cd} h_{ab,cd} + \Lambda^{ab,c} h_{ab,c} + \Lambda^{ab} h_{ab}. \tag{3.14}$$

† This quantity was first introduced by du Plessis [1].

If we substitute (3.11)–(3.13) in (3.14) we find

$$\begin{aligned}
JF &= \overline{\Lambda}^{ij,kh} \left[ \frac{\partial \bar{g}_{ij,kh}}{\partial g_{ab,cd}} h_{abcd} + \frac{\partial \bar{g}_{ij,kh}}{\partial g_{ab,c}} h_{ab,c} + \frac{\partial \bar{g}_{ij,kh}}{\partial g_{ab}} h_{ab} \right] \\
&+ \overline{\Lambda}^{ij,k} \left[ \frac{\partial \bar{g}_{ij,k}}{\partial g_{ab,c}} h_{ab,c} + \frac{\partial \bar{g}_{ij,k}}{\partial g_{ab}} h_{ab} \right] + \overline{\Lambda}^{ij} \left[ \frac{\partial \bar{g}_{ij}}{\partial g_{ab}} h_{ab} \right]. \tag{3.15}
\end{aligned}$$

From (3.8)–(3.10) it is easily seen that the quantities in square brackets in (3.15) are, respectively, $\bar{h}_{ij,kh}$, $\bar{h}_{ij,k}$, and $\bar{h}_{ij}$, so that (3.15) reduces to

$$JF = \bar{F}; \tag{3.16}$$

that is, $F$ is a scalar density. We now wish to find quantities $\Pi^{ij}$, $\Pi^{ij,k}$ so that $F$ can be expressed in the form

$$F = \Lambda^{ij,kl} h_{ij|k|l} + \Pi^{ij,k} h_{ij|k} + \Pi^{ij} h_{ij}. \tag{3.17}$$

Here the vertical stroke again represents covariant differentiation relative to the Christoffel symbols $\gamma_b^{\ a}{}_c$ defined in terms of the tensor $g_{ij}$ by (7.2.2), so that

$$h_{ij|k} = h_{ij,k} - \gamma_i^{\ a}{}_k h_{aj} - \gamma_j^{\ a}{}_k h_{ia}, \tag{3.18}$$

and

$$\begin{aligned}
h_{ij|k|l} &= h_{ij,kl} - \gamma_i^{\ a}{}_k h_{aj,l} - \gamma_j^{\ a}{}_k h_{ia,l} - \gamma_{i\ k,l}^{\ a} h_{aj} - \gamma_{j\ k,l}^{\ a} h_{ia} \\
&- \gamma_{k\ l}^{\ b} h_{bj,k} + \gamma_k^{\ b}{}_l (\gamma_b^{\ c}{}_k h_{cj} + \gamma_j^{\ c}{}_k h_{bc}) - \gamma_{j\ l}^{\ b} h_{ib,k} + \gamma_j^{\ b}{}_l (\gamma_i^{\ c}{}_k h_{cb} + \gamma_b^{\ c}{}_k h_{ic}) \\
&- \gamma_k^{\ b}{}_l h_{ij,b} + \gamma_k^{\ b}{}_l (\gamma_i^{\ c}{}_b h_{cj} + \gamma_j^{\ c}{}_b h_{ic}). \tag{3.19}
\end{aligned}$$

We substitute (3.18) and (3.19) in (3.17) and identify the (symmetrized) coefficients of $h_{ij}$ and $h_{ij,k}$ with those in (3.14) to find that $\Pi^{ij,k}$ and $\Pi^{ij}$ should be defined by

$$\Pi^{ij,k} = \Lambda^{ij,k} + 2\gamma_{a\ l}^{\ i} \Lambda^{aj,kl} + 2\gamma_{a\ l}^{\ j} \Lambda^{ai,kl} + \gamma_b^{\ k}{}_l \Lambda^{ij,bl}, \tag{3.20}$$

and

$$\begin{aligned}
\Pi^{ij} &= \Lambda^{ij} + \gamma_{a\ k,l}^{\ i} \Lambda^{aj,kl} + \gamma_{a\ k,l}^{\ j} \Lambda^{ai,kl} - \gamma_{a\ l}^{\ b} \gamma_b^{\ i}{}_k \Lambda^{aj,kl} - \gamma_c^{\ b}{}_l \gamma_b^{\ j}{}_k \Lambda^{ci,kl} \\
&- \gamma_b^{\ i}{}_l \gamma_c^{\ j}{}_k \Lambda^{bc,kl} - \gamma_b^{\ j}{}_l \gamma_c^{\ i}{}_k \Lambda^{bc,kl} - \gamma_k^{\ i}{}_l \gamma_c^{\ j}{}_b \Lambda^{cj,kl} - \gamma_k^{\ b}{}_l \gamma_c^{\ j}{}_b \Lambda^{ci,kl} \\
&+ \gamma_a^{\ i}{}_k \Pi^{aj,k} + \gamma_a^{\ j}{}_k \Pi^{ia,k}. \tag{3.21}
\end{aligned}$$

We now establish that $\Pi^{ij,k}$ and $\Pi^{ij}$ are tensorial. According to (3.11) and (3.16) the quantity $F - \Lambda^{ij,kl} h_{ij|k|l}$ is a scalar density, and hence it follows from (3.17) that $h_{ij} \Pi^{ij} + h_{ij|k} \Pi^{ij,k}$ is a scalar density for arbitrary symmetric type (0, 2) tensors $h_{ij}$. By an obvious generalization of the quotient theorems of Section 3.2, it follows that $(\Pi^{ij} + \Pi^{ji})$ and $(\Pi^{ij,k} + \Pi^{ji,k})$ are the components of tensor densities of type (2, 0) and (3, 0), respectively. However, from

(3.20) and (3.21)

$$\Pi^{ij} = \Pi^{ji}, \qquad \Pi^{ij,k} = \Pi^{ji,k}, \tag{3.22}$$

so that $\Pi^{ij}$ and $\Pi^{ij,k}$ are the components of tensor densities of type (2, 0) and (3, 0), respectively. These are the quantities which we were seeking.

Surprisingly enough, the relations (3.14) and (3.17) also yield an important additional result if we rewrite them in a more appropriate form. Clearly we have

$$h_{ij,k}\Lambda^{ij,k} = (h_{ij}\Lambda^{ij,k})_{,k} - h_{ij}\Lambda^{ij,k}{}_{,k}$$

and

$$\begin{aligned} h_{ij,kl}\Lambda^{ij,kl} &= (h_{ij,k}\Lambda^{ij,kl})_{,l} - h_{ij,k}\Lambda^{ij,kl}{}_{,l} \\ &= (h_{ij,l}\Lambda^{ij,kl})_{,k} - (h_{ij}\Lambda^{ij,kl}{}_{,l})_{,k} + h_{ij}\Lambda^{ij,kl}{}_{,lk}, \end{aligned}$$

so that (3.14) can be expressed in the form

$$F = -h_{ij}E^{ij}(L) + [h_{ij}\Lambda^{ij,k} + h_{ij,l}\Lambda^{ij,kl} - h_{ij}\Lambda^{ij,kl}{}_{,l}]_{,k} \tag{3.23}$$

where $E^{ij}(L)$ is the Euler–Lagrange operator (1.2) for the case (3.2), namely,

$$E^{ij}(L) = \frac{d}{dx^h}\left[\Lambda^{ij,h} - \frac{d}{dx^k}(\Lambda^{ij,hk})\right] - \Lambda^{ij}. \tag{3.24}$$

It is easily seen that (3.17) can be expressed in a similar manner, namely, as

$$F = -h_{ij}\{-\Lambda^{ij,kl}{}_{|k|l} + \Pi^{ij,k}{}_{|k} - \Pi^{ij}\} + [h_{ij}\Pi^{ij,k} + h_{ij|l}\Lambda^{ij,kl} - h_{ij}\Lambda^{ij,kl}{}_{|l}]_{|k}. \tag{3.25}$$

We observe that the quantities in square brackets in (3.25) are the components of a type (1, 0) tensor density, in which case the covariant derivative with respect to $x^k$ can be replaced by the corresponding partial derivative [see (4.1.27)]. We shall now establish the equivalence of the quantities which are respectively in square brackets in (3.23) and (3.25). From (3.20) we have

$$\begin{aligned} h_{ij}\Pi^{ij,k} + h_{ij|l}\Lambda^{ij,kl} - h_{ij}\Lambda^{ij,kl}{}_{|l} &= h_{ij}[\Lambda^{ij,k} + 4\gamma_a{}^i{}_l\Lambda^{aj,kl} + \gamma_b{}^k{}_l\Lambda^{ij,bl}] \\ &\quad + (h_{ij,l}\Lambda^{ij,kl} - 2\gamma_i{}^a{}_l h_{aj}\Lambda^{ij,kl}) \\ &\quad - h_{ij}(\Lambda^{ij,kl}{}_{,l} + 2\gamma_a{}^i{}_l\Lambda^{aj,kl} + \gamma_a{}^k{}_l\Lambda^{ij,al}) \\ &= h_{ij}\Lambda^{ij,k} + h_{ij,l}\Lambda^{ij,kl} - h_{ij}\Lambda^{ij,kl}{}_{,l}. \end{aligned}$$

Consequently (3.23) and (3.25) imply that

$$h_{ij}[E^{ij}(L) - \{-\Lambda^{ij,kl}{}_{|k|l} + \Pi^{ij,k}{}_{|k} - \Pi^{ij}\}] = 0$$

for arbitrary symmetric $h_{ij}$. This establishes the identity

$$E^{ij}(L) = -\Pi^{ij} + \Pi^{ij,k}{}_{|k} - \Lambda^{ij,kl}{}_{|k|l}, \tag{3.26}$$

which, in turn, demonstrates quite clearly that the $E^{ij}(L)$ are the components of a symmetric tensor density of type (2, 0).

We now wish to obtain the invariance identities associated with (3.2). These are a direct consequence of the fact that (3.3) is an identity in $J_{s\phantom{r}tu}^{\phantom{s}r}$, $J_{s\phantom{r}t}^{\phantom{s}r}$, and $J_s^r$ after the substitution of (3.8)–(3.10). Differentiating (3.3) with respect to $J_{s\phantom{r}tu}^{\phantom{s}r}$, noting (3.10), and taking into account the appropriate symmetry relations, we obtain

$$\overline{\Lambda}^{ij,kh}J_j^b\delta_r^a(\delta_i^s\delta_k^t\delta_h^u + \delta_i^s\delta_k^u\delta_h^t + \delta_i^t\delta_k^u\delta_h^s + \delta_i^t\delta_k^s\delta_h^u + \delta_i^u\delta_k^s\delta_h^t + \delta_i^u\delta_k^t\delta_h^s) = 0.$$

This is valid for an arbitrary transformation; thus, in particular, also for the identity transformation

$$\bar{x}^i = x^i \quad \text{with } J_i^j = \delta_i^j, \tag{3.27}$$

in which case it reduces to

$$\Lambda^{sb,tu} + \Lambda^{tb,us} + \Lambda^{ub,st} = 0 \tag{3.28}$$

by virtue of (3.7). This is a tensorial condition. By repeated application of (3.28) and (3.7), an important identity can be obtained as follows:

$$\Lambda^{sb,tu} = -\Lambda^{tb,us} - \Lambda^{ub,st} = \Lambda^{ts,bu} + \Lambda^{tu,sb} + \Lambda^{ut,bs} + \Lambda^{us,tb} = 2\Lambda^{tu,sb} - \Lambda^{bs,ut},$$

or

$$\Lambda^{sb,tu} = \Lambda^{tu,sb}. \tag{3.29}$$

We now return to (3.3) and differentiate it with respect to $J_{s\phantom{r}t}^{\phantom{s}r}$ to find, by (3.9) and (3.10),

$$\begin{aligned} \overline{\Lambda}^{ij,kh}\{[4J_j^bJ_h^c\delta_r^a(\delta_i^s\delta_k^t + \delta_k^s\delta_i^t) + J_j^aJ_j^b\delta_r^c(\delta_k^s\delta_h^t + \delta_h^s\delta_k^t)]g_{ab,c} \\ + 4\delta_r^a(\delta_i^s\delta_k^t + \delta_k^s\delta_i^t)J_j^b{}_h g_{ab}\} + 2\overline{\Lambda}^{ij,k}\delta_r^a(\delta_i^s\delta_k^t + \delta_i^t\delta_k^s)J_j^b g_{ab} = 0. \end{aligned}$$

For the special transformation (3.27) this reduces to

$$2\Lambda^{sb,tc}g_{rb,c} + 2\Lambda^{tb,sc}g_{rb,c} + \Lambda^{ab,st}g_{ab,r} + \Lambda^{sb,t}g_{rb} + \Lambda^{tb,s}g_{rb} = 0. \tag{3.30}$$

We now wish to express (3.30) in a more useful form, and this we do by introducing a normal coordinate system, as discussed in Section 4.3, where it should be remarked that this is permissible by virtue of the fact that the Christoffel symbols define a symmetric connection on our manifold $V_n$. At the pole $P$ of this normal coordinate system, where $P$ is chosen arbitrarily, we have $\gamma_j{}^i{}_k = 0$, so that by (7.2.3), $g^{ij}{}_{,k} = 0$ and $g_{ij,k} = 0$. A comparison of (3.20) and (3.30) shows that

$$\Pi^{sb,t}g_{rb} + \Pi^{tb,s}g_{rb} = 0,$$

that is,

$$\Pi^{su,t} + \Pi^{tu,s} = 0. \tag{3.31}$$

From (3.22) and (3.31) we thus find

$$\Pi^{su,t} = -\Pi^{tu,s} = \Pi^{ts,u} = -\Pi^{us,t},$$

or

$$\Pi^{su,t} = 0, \tag{3.32}$$

at $P$. However, since $\Pi^{su,t}$ is tensorial, this is valid in an arbitrary coordinate system, and furthermore, since $P$ is arbitrary, the relation (3.32) is valid everywhere. Thus (3.32) is an identity which holds throughout $V_n$. It is, in fact, equivalent to (3.30).

*Remark 1.* From (3.32) it now becomes evident why we have concentrated our attention on scalar densities of the type (3.2) rather than on the more natural (and simpler) choice

$$L = L(g_{ij}, g_{ij,k}).$$

In the latter case $\Lambda^{ij,hk} = 0$, which implies by (3.20) that

$$\Pi^{ij,k} = \Lambda^{ij,k}.$$

However, the identity (3.32) then ensures that $\partial L / \partial g_{ij,k} = 0$. Thus we have the following theorem.

**THEOREM**

*There does not exist a scalar density $L = L(g_{ij}, g_{ij,k})$ which depends solely on the $g_{ij}$ and their first derivatives.*

*Remark 2.* Instead of confining our attention to scalar densities of the type (3.2), we may consider relative tensors of type $(r, s)$ and weight $w$, with the same functional dependence as in (3.2), namely,

$$A^{i_1 \cdots i_r}_{j_1 \cdots j_s} = A^{i_1 \cdots i_r}_{j_1 \cdots j_s}(g_{ij}, g_{ij,k}, g_{ij,kh}),$$

and define

$$A^{i_1 \cdots i_r; ab,cd}_{j_1 \cdots j_s} = \frac{\partial A^{i_1 \cdots i_r}_{j_1 \cdots j_s}}{\partial g_{ab,cd}}, \qquad A^{i_1 \cdots i_r; ab,c}_{j_1 \cdots j_s} = \frac{\partial A^{i_1 \cdots i_r}_{j_1 \cdots j_s}}{\partial g_{ab,c}}.$$

The analysis giving rise to (3.11), (3.28), and (3.32) remains essentially unchanged. The counterparts of the latter imply that $A^{i_1 \cdots i_r; ab,cd}_{j_1 \cdots j_s}$ are the components of a relative tensor of type $(r + 4, s)$ and weight $w$, which satisfies

$$A^{i_1 \cdots i_r; ab,cd}_{j_1 \cdots j_s} + A^{i_1 \cdots i_r; db,ac}_{j_1 \cdots j_s} + A^{i_1 \cdots i_r; cb,da}_{j_1 \cdots j_s} = 0,$$

while

$$A^{i_1 \cdots i_r; ab,c}_{j_1 \cdots j_s} + \gamma_{h\ k}^{\ c} A^{i_1 \cdots i_r; ab,hk}_{j_1 \cdots j_s} + 2\gamma_{h\ k}^{\ a} A^{i_1 \cdots i_r; hb,ck}_{j_1 \cdots j_s} + 2\gamma_{h\ k}^{\ b} A^{i_1 \cdots i_r; ah,ck}_{j_1 \cdots j_s} = 0.$$

Again we see that if $A^{i_1 \cdots i_r; ab,cd}_{j_1 \cdots j_s} = 0$, then $A^{i_1 \cdots i_r; ab,c}_{j_1 \cdots j_s} = 0$; that is, there are no relative tensors the components of which depend on $g_{ab,c}$ explicitly (see also McKiernan and Richards [1]).

We now return to (3.3) once more, and differentiate it with respect to $J_s^r$, which, by (2.13) and (3.8)–(3.10), yields

$$\bar{\Lambda}^{ij,kh}(\delta_r^a \delta_i^s J_k^b J_h^d + J_i^a \delta_r^b \delta_j^s J_k^c J_h^d + J_i^a J_j^b \delta_r^c \delta_k^s J_h^d + J_i^a J_j^b J_k^c \delta_r^d \delta_h^s)g_{ab,cd}$$
$$+ \bar{\Lambda}^{ij,k}(\delta_r^a \delta_i^s J_j^b J_k^c + J_i^a \delta_r^b \delta_j^s J_k^c + J_i^a J_j^b \delta_r^c \delta_k^s)g_{ab,c}$$
$$+ \bar{\Lambda}^{ij}(\delta_r^a \delta_i^s J_j^b + J_i^a \delta_r^b \delta_j^s)g_{ab} + \mu_r^s = J A_s^r L, \tag{3.33}$$

where $\mu_r^s$ contains terms which are linear in $J_b^{\ a}{}_c$ and $J_b^{\ a}{}_{cd}$. For the special transformation (3.27) we see that $\mu_r^s = 0$, in which case (3.33) reduces to (Buchdahl [1])

$$2\Lambda^{sb,cd}g_{rb,cd} + 2\Lambda^{ab,sd}g_{ab,rd} + 2\Lambda^{sb,c}g_{rb,c} + \Lambda^{ab,s}g_{ab,r} + 2\Lambda^{sb}g_{rb} = \delta_r^s L. \tag{3.34}$$

We now wish to rewrite (3.34) in a more useful form. In doing so we will frequently make use of the following result, which follows from (3.7) and (3.28): if $\alpha_{ij}$ is any quantity which is symmetric in $i$ and $j$, then

$$\alpha_{ij}\Lambda^{ai,bj} = \tfrac{1}{2}\alpha_{ij}(\Lambda^{ai,bj} + \Lambda^{aj,bi}) = -\tfrac{1}{2}\alpha_{ij}\Lambda^{ab,ij}. \tag{3.35}$$

By (7.3.16), at the pole $P$ of a normal coordinate system, we have

$$R_{brad} = \tfrac{1}{2}(g_{bd,ra} + g_{ra,bd} - g_{dr,ab} - g_{ab,dr}),$$

so that a repeated application of (3.35) yields

$$\Lambda^{ab,sd}R_{brad} = -\tfrac{3}{4}\Lambda^{ab,sd}(g_{ab,dr} + g_{dr,ab}) \tag{3.36}$$

at $P$. Furthermore, by (3.21) and (3.32)

$$\Lambda^{ts} = \Pi^{ts} - \gamma_{a\ k,l}^{\ t}\Lambda^{as,kl} - \gamma_{a\ k,l}^{\ s}\Lambda^{at,kl}. \tag{3.37}$$

Since at $P$

$$\gamma_{a\ k,l}^{\ t} = \tfrac{1}{2}g^{tb}(g_{ab,kl} + g_{kb,al} - g_{ak,bl}), \tag{3.38}$$

the relation (3.37) can be expressed, by repeated application of (3.35), in the form

$$\Lambda^{ts} = \Pi^{ts} - \tfrac{1}{4}g^{th}(g_{ah,kl} + g_{lk,ah})\Lambda^{as,kl} - \tfrac{1}{4}g^{sh}(g_{ah,kl} + g_{lk,ah})\Lambda^{at,kl} \tag{3.39}$$

at $P$. If we apply (3.36) twice to (3.39), we then obtain

$$\Lambda^{ts} = \Pi^{ts} + \tfrac{1}{3}g^{th}\Lambda^{ab,sd}R_{bhad} + \tfrac{1}{3}g^{sh}\Lambda^{ab,td}R_{bhad}. \tag{3.40}$$

We now substitute (3.36) and (3.40) in (3.34) to find, at the pole $P$,

$$-\Lambda^{ab,sd}R_{b\ ad}^{\ r} + \Pi^{rs} + \tfrac{1}{3}\Lambda^{ab,rd}R_{b\ ad}^{\ s} = \tfrac{1}{2}g^{rs}L. \tag{3.41}$$

Using similar arguments to those following (3.32), we conclude that (3.41) is valid everywhere in all coordinate systems. In view of (3.22) the relation (3.41) thus gives rise to

$$\Pi^{rs} = \tfrac{1}{2}g^{rs}L + \tfrac{2}{3}\Lambda^{ab,\,sd}R_b{}^r{}_{ad}, \tag{3.42}$$

which is, in fact, equivalent to (3.34). Thus for the Lagrangian (3.2) the *invariance identities* are (3.28), (3.32), and (3.42).

If we return to (3.26), and note (3.32) and (3.42), we find

$$E^{ij}(L) = -\tfrac{1}{2}g^{ij}L - \tfrac{2}{3}\Lambda^{ab,\,id}R_b{}^j{}_{ad} - \Lambda^{ij,\,kl}{}_{|k|l}. \tag{3.43}$$

This form of the Euler-Lagrange expression is particularly useful, since it indicates that it is only necessary to evaluate $\Lambda^{ij,\,kl}$ when an explicit expression for $E^{ij}(L)$ is sought.

Finally, we turn to the evaluation of $E^{ij}(L)_{|j}$ which we again evaluate at the pole $P$ of a normal coordinate system. From (3.24) we have

$$E^{ij}(L)_{|j} = -\Lambda^{ij,\,hk}{}_{,hkj} + \Lambda^{ij,\,h}{}_{,hj} - \Lambda^{ij}{}_{,j}. \tag{3.44}$$

By virtue of (3.28) the first term on the right-hand side vanishes identically, so we need consider only the remaining two. In view of (3.20), (3.32), and (3.35) we have

$$\Lambda^{ij,\,k} = \gamma_a{}^i{}_l \Lambda^{al,\,kj} + \gamma_a{}^j{}_l \Lambda^{al,\,ki} - \gamma_a{}^k{}_l \Lambda^{al,\,ij},$$

so that

$$\Lambda^{ij,\,k}{}_{,kj} = (\gamma_a{}^i{}_l \Lambda^{al,\,kj})_{,kj},$$

which, at the pole, reduces to

$$\Lambda^{ij,\,k}{}_{,kj} = \gamma_a{}^i{}_{l,\,kj}\Lambda^{al,\,kj} + 2\gamma_a{}^i{}_{l,\,k}\Lambda^{al,\,kj}{}_{,j}. \tag{3.45}$$

In view of (3.38) and

$$\gamma_a{}^i{}_{l,\,kj} = \tfrac{1}{2}g^{ih}(g_{ah,\,lkj} + g_{lh,\,akj} - g_{al,\,hkj})$$

at $P$, we see that (3.45) gives rise to

$$\Lambda^{ij,\,k}{}_{,kj} = -\tfrac{1}{2}g^{ih}g_{al,\,hkj}\Lambda^{al,\,kj} - g^{ih}(g_{al,\,hk} + g_{kh,\,al})\Lambda^{al,\,jk}{}_{,j}, \tag{3.46}$$

where we have made use of (3.28) and (3.35).

However, from (3.34), we have

$$\Lambda^{ij} = \tfrac{1}{2}g^{ij}L - g_{lb,\,c}g^{li}\Lambda^{jb,\,c} - \tfrac{1}{2}\Lambda^{ab,\,j}g_{ab,\,l}g^{li} - g^{il}\Lambda^{jd,\,ab}(g_{ld,\,ab} + g_{ab,\,ld}),$$

so that at $P$,

$$\Lambda^{ij}{}_{,j} = \tfrac{1}{2}g^{ij}\Lambda^{ab,\,cd}g_{ab,\,cdj} - g^{il}\Lambda^{jd,\,ab}{}_{,j}(g_{ld,\,ab} + g_{ab,\,ld}) - g^{il}\Lambda^{jd,\,ab}g_{ab,\,ldj}. \tag{3.47}$$

If we substitute (3.46) and (3.47) in (3.44), noting (3.29), we find, at $P$

$$E^{ij}(L)_{|j} = 0. \tag{3.48}$$

As before this is a tensorial identity and is valid throughout $V_n$ in any coordinate system. This establishes the following theorem.

**THEOREM**

*For any scalar density $L = L(g_{ij}, g_{ij,\,h}, g_{ij,\,hk})$ the components $E^{ij}(L)$ of the Euler-Lagrange tensor density are such that their divergence vanishes identically.*

We shall conclude this section by looking at the specific scalar density defined by

$$L_0 = \alpha\sqrt{g}\,R - 2\lambda\sqrt{g} = \alpha\sqrt{g}\,g^{ac}g^{bd}R_{abcd} - 2\lambda\sqrt{g}, \tag{3.49}$$

where $\alpha$, $\lambda$ are constants and $R$ and $R_{abcd}$ are defined by (7.3.26) and (7.3.16). Clearly $L_0$ is a scalar density of the type (3.2) so that we may apply the above analysis in order to calculate $E^{ij}(L_0)$. As indicated earlier, this requires only a knowledge of $\Lambda^{ij,\,kl}$ for $L_0$, which we now evaluate. From (7.3.16) and the remark following (3.4) we see that

$$\frac{\partial R_{abcd}}{\partial g_{ij,\,kl}} = \tfrac{1}{8}[(\delta_a^i\delta_d^j + \delta_d^i\delta_a^j)(\delta_b^k\delta_c^l + \delta_c^k\delta_b^l) + (\delta_b^i\delta_c^j + \delta_c^i\delta_b^j)(\delta_a^k\delta_d^l + \delta_d^k\delta_a^l)$$
$$-(\delta_a^i\delta_c^j + \delta_c^i\delta_a^j)(\delta_b^k\delta_d^l + \delta_d^k\delta_b^l) - (\delta_b^i\delta_d^j + \delta_d^i\delta_b^j)(\delta_a^k\delta_c^l + \delta_c^k\delta_a^l)]. \tag{3.50}$$

Consequently, from (3.49) and (3.50), we have

$$\Lambda^{ij,\,kl} = -\alpha\sqrt{g}\,\frac{2g^{ij}g^{kl} - g^{ik}g^{jl} - g^{il}g^{jk}}{2}, \tag{3.51}$$

which, when substituted in (3.43), yields by virtue of Ricci's lemma,

$$E^{hk}(L_0) = \alpha\sqrt{g}\,[R^{hk} - \tfrac{1}{2}g^{hk}R] + \lambda\sqrt{g}\,g^{hk}, \tag{3.52}$$

where $R^{hk}$ is defined by (7.3.29). We note that the quantities in square brackets are the components of the Einstein tensor defined by (7.3.30). We see that the Euler-Lagrange equations

$$E^{ij}(L) = 0, \tag{3.53}$$

in the case (3.49), reduce to

$$\alpha(R^{ij} - \tfrac{1}{2}g^{ij}R) + \lambda g^{ij} = 0. \tag{3.54}$$

*These equations are the well-known Einstein vacuum field equations* (with cosmological term) provided that the $g_{hk}$ are interpreted as the components of the metric tensor of four-dimensional space-time.†

*Remark*   In general, it is clear from (3.24) that $E^{ij}(L)$ will be of fourth order in $g_{ab}$. However, for the particular scalar density $L_0$, given by (3.49), the associated $E^{ij}(L_0)$ as given by (3.52) are only of second order. One might be tempted to think that (3.54) could be obtained from a scalar density of the type $L = L(g_{ij}, g_{ij,k})$ since the associated Euler-Lagrange expression would be of second order. However, we have already seen that such scalar densities do not exist.

## 8.4   THE FIELD EQUATIONS OF EINSTEIN IN VACUO

In the preceding section we saw that the Einstein vacuum field equations (3.54) are obtainable from a variational principle when the scalar density $L$ is chosen suitably, namely, as given by (3.49). In view of the physical significance attributed to these equations in the four-dimensional case, it seems appropriate to consider their derivation in a little more detail. From (3.26) and (3.48) we see that in general the $E^{ij}(L)$ are the components of a type $(2, 0)$ tensor density whose functional dependence is given by

$$E^{ij}(L) = E^{ij}(g_{ab}, g_{ab,r}, g_{ab,rs}, g_{ab,rst}, g_{ab,rstu}),$$

and which, in addition, satisfy the conditions

$$E^{ij}(L)_{|j} = 0, \qquad E^{ij}(L) = E^{ji}(L).$$

However, if $L$ is chosen according to (3.49) then the $E^{ij}(L)$ are merely dependent on the $g_{ab}$ and its first and second derivatives. Consequently, guided by these observations, it seems natural to consider the following problem: to find, in a four-dimensional space, all type $(2, 0)$ tensor densities the components $A^{ij}$ of which satisfy the conditions (Lovelock [3, 5])

$$A^{ij} = A^{ij}(g_{ab}, g_{ab,c}, g_{ab,cd}), \tag{4.1}$$

$$A^{ij}_{\ |j} = 0 \tag{4.2}$$

and

$$A^{ij} = A^{ji}. \tag{4.3}$$

With regard to this problem it should be noted that, in the first instance, it is *not* assumed that $A^{ij}$ represent an Euler-Lagrange tensor density $E^{ij}(L)$.

† For an introduction to the general theory of relativity see the following texts: Adler, Bazin, and Schiffer [1], Eddington [1], Fock [1], Hawking and Ellis [1], Møller [1], Pauli [1], Schrödinger [1], Synge [4], and Weyl [2].

If we introduce the notation

$$\left.\begin{aligned} A^{ij;ab,cd} &= \frac{\partial A^{ij}}{\partial g_{ab,cd}}, \\[2mm] A^{ij;ab,c} &= \frac{\partial A^{ij}}{\partial g_{ab,c}}, \\[2mm] A^{ij;ab} &= \frac{\partial A^{ij}}{\partial g_{ab}}, \end{aligned}\right\} \tag{4.4}$$

and recall that $A^{ij}$ are assumed to be the components of a tensor density, then the analysis giving rise to (3.11)–(3.13) establishes that only $A^{ij;ab,cd}$ are the components of a tensor density. Furthermore, the counterpart of the invariance identity (3.28) is easily shown to be

$$A^{ij;ab,cd} + A^{ij;db,ac} + A^{ij;cb,da} = 0, \tag{4.5}$$

which, in view of the obvious properties

$$A^{ij;ab,cd} = A^{ij;ba,cd} = A^{ij;ab,dc}, \tag{4.6}$$

implies that

$$A^{ij;ab,cd} = A^{ij;cd,ab} \tag{4.7}$$

[see (3.29)].

Because for a type $(2, 0)$ tensor density we have

$$A^{ij}_{\ |j} = \frac{\partial A^{ij}}{\partial x^j} + \gamma_{h\ j}^{\ i}A^{hj} + \gamma_{h\ j}^{\ j}A^{ih} - \gamma_{j\ h}^{\ h}A^{ij},$$

it follows from (4.1) and (4.4) that

$$A^{ij}_{\ |j} = A^{ij;ab,cd}g_{ab,cdj} + A^{ij;ab,c}g_{ab,cj} + A^{ij;ab}g_{ab,j} + \gamma_{h\ j}^{\ i}A^{hj},$$

from which it is inferred that

$$\frac{\partial(A^{ij}_{\ |j})}{\partial g_{ab,rst}} = \frac{1}{3}\left(A^{it;ab,rs} + A^{is;ab,tr} + A^{ir;ab,st}\right).$$

Consequently, if (4.2) is applied to the latter we find

$$A^{it;ab,rs} + A^{is;ab,tr} + A^{ir;ab,st} = 0. \tag{4.8}$$

However, by (4.3) and (4.4) we also have

$$A^{it;ab,rs} = A^{ti;ab,rs}, \tag{4.9}$$

so that (4.8) and (4.9) imply [see (3.29)]

$$A^{it;ab,rs} = A^{rs;ab,it}. \tag{4.10}$$

From (4.7) and (4.10) we thus have

$$A^{it;\,ab,\,rs} = A^{ab;\,it,\,rs}, \tag{4.11}$$

which, by (4.8), establishes the relation

$$A^{it;\,ab,\,rs} + A^{ib;\,ta,\,rs} + A^{ia;\,bt,\,rs} = 0. \tag{4.12}$$

If we now introduce the quantities

$$A^{ij;\,ab,\,cd;\,rs,\,tu} = \frac{\partial A^{ij;\,ab,\,cd}}{\partial g_{rs,\,tu}}, \tag{4.13}$$

then it is easily seen that they form the components of a tensor density (see Section 8.3, Remark 2). Obviously, by definition, $A^{ij;\,ab,\,cd;\,rs,\,tu}$ has the same symmetry properties in the indices $ijabcd$ as does $A^{ij;\,ab,\,cd}$, namely, (4.5)–(4.12). Furthermore, from (4.4) and (4.13),

$$A^{ij;\,ab,\,cd;\,rs,\,tu} = A^{ij;\,rs,\,tu;\,ab,\,cd}. \tag{4.14}$$

Consequently, the quantities $A^{i_1 i_2;\,i_3 i_4,\,i_5 i_6;\,i_7 i_8,\,i_9 i_{10}}$ are the components of a tensor density with the following symmetry properties:

1. It is symmetric in $i_{2h-1} i_{2h}$ for $h = 1, \ldots, 5$.
2. It is symmetric under interchange of the pair $(i_{2h-1}, i_{2h})$ with the pair $(i_{2k-1}, i_{2k})$ for $h, k = 1, \ldots, 5$.
3. It satisfies a cyclic identity [as characterized by (4.12)] involving any three of the four indices $(i_{2h-1}, i_{2h})$ $(i_{2k-1}, i_{2k})$ for $h, k = 1, \ldots, 5$ $(h \neq k)$.

We now wish to show that in a four-dimensional space

$$A^{ij;\,ab,\,cd;\,rs,\,tu} = 0 \qquad (n = 4). \tag{4.15}$$

This is accomplished in the following manner. In view of the fact that (4.13) has 10 free indices, at least three of these must coincide in a four-dimensional space. Let this index be denoted by 1. By properties 1 and 2 these three coinciding indices can always be brought into one of the two following situations: $A^{i1;\,1b,\,1d;\,rs,\,tu}$ or $A^{i1;\,11,\,bd;\,rs,\,tu}$. However, the second of these vanishes identically by an application of property 3 to the second, third, and fourth indices. Furthermore, if we use property 3 on the third, fourth, and fifth indices of $A^{i1;\,1b,\,1d;\,rs,\,tu}$ we find

$$2A^{i1;\,1b,\,1d;\,rs,\,tu} = -A^{i1;\,11,\,bd;\,rs,\,tu},$$

and we have already seen that the right-hand side of the latter vanishes. This establishes (4.15). For future reference we draw attention to the fact that this is the only occasion on which we make use of the assumption that $n = 4$.

In order to obtain $A^{ij}$ we integrate the differential equations (4.15). From (4.13) and (4.15) it therefore follows that

$$A^{ij;\,ab,\,cd} = \alpha^{ijabcd}, \tag{4.16}$$

where $\alpha^{ijabcd}$ are the components of a tensor density which has all the symmetry properties of $A^{ij;\,ab,\,cd}$, namely, (4.5)–(4.12), and which is at most a function of $g_{rs}$ and $g_{rs,\,t}$. However, we have already indicated that all tensorial quantities of this type must be independent of $g_{rs,\,t}$ (see Section 8.3, Remark 2), so that

$$\alpha^{ijabcd} = \alpha^{ijabcd}(g_{rs}). \tag{4.17}$$

From (4.4), (4.16), and (4.17) we thus have

$$A^{ij} = \alpha^{ijabcd} g_{ab,\,cd} + \beta^{ij}(g_{rs}, g_{rs,\,t}), \tag{4.18}$$

where $\beta^{ij}$ are, as yet, undetermined functions. We now wish to express the right-hand side of (4.18) in tensorial form. Now from (7.3.16) we have

$$R_{bcad} = \tfrac{1}{2}(g_{bd,\,ca} + g_{ca,\,bd} - g_{dc,\,ab} - g_{ab,\,dc}) + \lambda_{bcad},$$

where

$$\lambda_{bcad} = \gamma_{cma}\gamma_b{}^m{}_d - \gamma_{cmd}\gamma_a{}^m{}_b. \tag{4.19}$$

Therefore, by the technique exemplified by (3.35) applied twice, we find

$$\begin{aligned}
\alpha^{ijabcd} R_{bcad} &= -\tfrac{3}{4}\alpha^{ijabcd}(g_{ab,\,cd} + g_{dc,\,ab}) + \alpha^{ijabcd}\lambda_{bcad} \\
&= -\tfrac{3}{2}\alpha^{ijabcd} g_{ab,\,cd} + \alpha^{ijabcd}\lambda_{bcad},
\end{aligned}$$

where we have used (4.7) in the last step. From the latter equation and (4.18) we thus find

$$A^{ij} = -\tfrac{2}{3}\alpha^{ijabcd} R_{bcad} + \mu^{ij} \tag{4.20}$$

where

$$\mu^{ij} = \beta^{ij} + \tfrac{2}{3}\alpha^{ijabcd}\lambda_{bcad}.$$

However, in (4.20) the quantities $A^{ij} + \tfrac{2}{3}\alpha^{ijabcd} R_{bcad}$ are the components of a symmetric type (2, 0) tensor density, in which case, so are $\mu^{ij}$. Moreover, from (4.18) and (4.19), the $\mu^{ij}$ are independent of $g_{ab,\,cd}$, which implies that they are also independent of $g_{ab,\,c}$, that is,

$$\mu^{ij} = \mu^{ij}(g_{ab}).$$

Our problem has thus reduced to the evaluation of $\mu^{ij}$ and $\alpha^{ijabcd}$. We shall first find $\mu^{ij}$, the most general symmetric type (2, 0) tensor density which is a function of $g_{ab}$. Since $\mu^{ij}$ is a tensor density, its transformation law

is explicitly given by

$$J_j^k J_i^l \bar{\mu}^{ij}(J_s^r J_a^c g_{rc}) = J\mu^{lk}, \tag{4.21}$$

where $J_j^k$ and $J$ are defined by (1.7) and (1.6). By differentiating (4.21) with respect to $J_b^a$, and subsequently setting $\bar{x}^i = x^i$, we obtain the invariance identity

$$(\delta_a^k \delta_j^b \delta_i^l + \delta_j^k \delta_a^l \delta_i^b)\mu^{ij} + \delta_j^k \delta_i^l(\delta_a^r \delta_s^b \delta_d^c + \delta_s^r \delta_a^c \delta_d^b)g_{rc}\frac{\partial \mu^{ij}}{\partial g_{sd}} = \delta_a^b \mu^{lk},$$

where use has been made of (2.13). This equation reduces to

$$\delta_a^k \mu^{lb} + \delta_a^l \mu^{bk} + 2g_{ac}\frac{\partial \mu^{lk}}{\partial g_{bc}} = \delta_a^b \mu^{lk},$$

which upon multiplication by $g^{ad}$ can be written in the form

$$g^{kd}\mu^{lb} + g^{ld}\mu^{bk} = g^{bd}\mu^{lk} - 2\frac{\partial \mu^{lk}}{\partial g_{bd}}. \tag{4.22}$$

The right-hand side of (4.22) is symmetric in $b$ and $d$, so that we have

$$g^{kd}\mu^{lb} + g^{ld}\mu^{bk} = g^{kb}\mu^{ld} + g^{lb}\mu^{dk}.$$

If we multiply the latter by $g_{kd}$ and note that $\mu^{ij}$ is symmetric, we have

$$\mu^{lb} = \lambda g^{lb}, \tag{4.23}$$

where

$$\lambda = \frac{\mu^{ij}g_{ij}}{n}.$$

We now substitute (4.23) into (4.22), and note the identity

$$\frac{\partial g^{lk}}{\partial g_{bd}} = -\frac{1}{2}g^{lb}g^{dk} + g^{ld}g^{bk}$$

which follows from (7.1.34), to find

$$2\frac{\partial \lambda}{\partial g_{bd}} = \lambda g^{bd}. \tag{4.24}$$

If we define $\psi$ by

$$\lambda = \sqrt{g}\,\psi, \tag{4.25}$$

and recall that $\partial g/\partial g_{bd} = gg^{bd}$, we see that (4.24) reduces to $\partial \psi/\partial g_{bd} = 0$; that is, $\psi = a$, where $a$ is a constant. We now substitute this together with

(4.25) in (4.23), to find

$$\mu^{lb} = a\sqrt{g}\,g^{lb}. \tag{4.26}$$

We have thus established the following lemma.

**LEMMA**

*The most general symmetric type* (2, 0) *tensor density which is a function of* $g_{rs}$ *alone, has components* $a\sqrt{g}\,g^{ij}$, *where* $a$ *is an arbitrary constant.*

*Remark.*   This lemma is a special case of a more general result in which no symmetry properties of the tensor are assumed, namely, if $\alpha^{ij}$ are the components of a type (2, 0) tensor density and $\alpha^{ij} = \alpha^{ij}(g_{rs})$ then

$$\alpha^{ij} = \begin{cases} a\sqrt{g}\,g^{ij} & \text{for } n > 2, \\ a\sqrt{g}\,g^{ij} + b\varepsilon^{ij} & \text{for } n = 2, \end{cases}$$

where $a$ and $b$ are constants. We note that, for $n > 2$, the quantities $\alpha^{ij}$ are inevitably symmetric (Lovelock [2]).

We now turn to the calculation of $\alpha^{ijabcd}$ which follows the same pattern as the above, but is a little more complicated. The analysis by which (4.22) was obtained from (4.21) is now repeated for $\alpha^{ijabcd}$ and we find

$$g^{ir}\alpha^{sjabcd} + g^{jr}\alpha^{isabcd} + g^{ar}\alpha^{ijsbcd} + g^{br}\alpha^{ijascd} + g^{cr}\alpha^{ijabsd} + g^{dr}\alpha^{ijabcs}$$
$$= g^{rs}\alpha^{ijabcd} - 2\frac{\partial \alpha^{ijabcd}}{\partial g_{rs}}. \tag{4.27}$$

Again the right-hand side of (4.27) is symmetric in $r$ and $s$, so that we have

$$g^{ir}\alpha^{sjabcd} + g^{jr}\alpha^{isabcd} + g^{ar}\alpha^{ijsbcd} + g^{br}\alpha^{ijascd} + g^{cr}\alpha^{ijabsd} + g^{dr}\alpha^{ijabcs}$$
$$= g^{is}\alpha^{rjabcd} + g^{js}\alpha^{irabcd} + g^{as}\alpha^{ijrbcd} + g^{bs}\alpha^{ijarcd} + g^{cs}\alpha^{ijabrd} + g^{ds}\alpha^{ijabcr},$$

which, when multiplied by $g_{ir}$, gives

$$(n - 1)\alpha^{sjabcd} + \alpha^{isabcd} + \alpha^{ajsbcd} + \alpha^{bjascd} + \alpha^{cjabsd} + \alpha^{djabcs}$$
$$= g^{js}g_{ir}\alpha^{irabcd} + g^{as}g_{ir}\alpha^{ijrbcd} + g^{bs}g_{ir}\alpha^{ijarcd} + g^{cs}g_{ir}\alpha^{ijabrd} + g^{ds}g_{ir}\alpha^{ijabcr}. \tag{4.28}$$

By (4.9) the second term on the left-hand side can be combined with the first term. Furthermore, it is possible to combine the remaining terms on the left-hand side with the first term, since, by (4.12) and (4.8), we have

$$\alpha^{ajsbcd} + \alpha^{bjascd} = -\alpha^{sjbacd}$$

and

$$\alpha^{cjabsd} + \alpha^{djabcs} = -\alpha^{sjabdc}.$$

Consequently the left-hand side of (4.28) reduces to $(n-2)\alpha^{sjabcd}$. We now turn to the right-hand side and consider the coefficient of $g^{as}$. From (4.12) we have

$$\alpha^{ijrbcd} + \alpha^{rjbicd} = -\alpha^{bjircd},$$

which, when multiplied by $g_{ir}$, account being taken of (4.6), yields

$$g_{ir}\alpha^{ijrbcd} = -\frac{\phi^{bjcd}}{2},$$

where

$$\phi^{bjcd} = g_{ir}\alpha^{bjircd}. \tag{4.29}$$

This same process is repeated on the third, fourth, and fifth terms on the right-hand side of (4.28), so that the latter reduces to

$$2(n-2)\alpha^{sjabcd} = 2g^{js}\phi^{abcd} - g^{as}\phi^{bjcd} - g^{bs}\phi^{ajcd} - g^{cs}\phi^{abjd} - g^{ds}\phi^{abcj}. \tag{4.30}$$

Consequently (for $n > 2$) the problem of finding $\alpha^{sjabcd}$ has been reduced to that of calculating $\phi^{abjd}$. Furthermore, by (4.29), $\phi^{bjcd}$ are the components of a tensor density with the symmetry properties

$$\left.\begin{array}{l} \phi^{abjd} = \phi^{bajd} = \phi^{abdj} = \phi^{jdab}, \\ \phi^{abjd} + \phi^{adbj} + \phi^{ajdb} = 0. \end{array}\right\} \tag{4.31}$$

Moreover, since $\phi^{abjd}$ is a function of $g_{rs}$ alone, the analysis which gave rise to (4.27) can be repeated for $\phi^{abjd}$, and we find

$$g^{ir}\phi^{sjab} + g^{jr}\phi^{isab} + g^{ar}\phi^{ijsb} + g^{br}\phi^{ijas} = g^{rs}\phi^{ijab} - 2\frac{\partial \phi^{ijab}}{\partial g_{rs}}.$$

Again the symmetry of the right-hand side in $r$ and $s$ is used to obtain an equation similar to (4.28), namely,

$$(n-1)\phi^{sjab} + \phi^{jsab} + \phi^{ajsb} + \phi^{bjas} = g^{js}g_{ir}\phi^{irab} + g^{as}g_{ir}\phi^{ijrb} + g^{bs}g_{ir}\phi^{ijar},$$

which, by (4.31), can be written in a form similar to (4.30), namely,

$$(n-1)\phi^{sjab} = g^{js}\phi^{ab} - \tfrac{1}{2}g^{as}\phi^{jb} - \tfrac{1}{2}g^{bs}\phi^{ja}, \tag{4.32}$$

where

$$\phi^{ab} = g_{ir}\phi^{irab}.$$

Clearly, $\phi^{ab}$ are the components of a symmetric type $(2,0)$ tensor density which is a function of $g_{rs}$ only, which by the previous lemma implies that

$$\phi^{ab} = \beta\sqrt{g}\,g^{ab}, \tag{4.33}$$

where $\beta$ is a constant. After substituting (4.33) in (4.32), which in turn is substituted in (4.30), we have completely determined $\alpha^{ijabcd}$.

We now return to (4.20) in order to evaluate $A^{ij}$. From (4.30) and (7.3.11) we have

$$(n-2)\alpha^{sjabcd}R_{bcad} = g^{js}\phi^{abcd}R_{bcad} - 2\phi^{bjcd}R_{bcad}g^{as}. \tag{4.34}$$

However, from (4.32) and (4.33) we see that

$$(n-1)\phi^{bjcd}R_{bcad} = \beta\sqrt{g}\,(g^{jb}g^{cd} - \tfrac{1}{2}g^{bc}g^{jd} - \tfrac{1}{2}g^{bd}g^{cj})R_{bcad}$$
$$= \frac{3\beta\sqrt{g}\,R_a^j}{2},$$

where $R_a^j$ is defined by (7.3.25). Consequently for $n > 2$, the relation (4.34) becomes

$$\alpha^{sjabcd}R_{bcad} = \frac{3\beta\sqrt{g}\,[\tfrac{1}{2}g^{js}R - R^{js}]}{(n-1)(n-2)},$$

where $R$ is defined by (7.3.26). The latter, together with (4.26), is substituted in (4.20) to finally yield

$$A^{ij} = \alpha\sqrt{g}\,[R^{ij} - \tfrac{1}{2}g^{ij}R] + \lambda\sqrt{g}\,g^{ij}. \tag{4.35}$$

We have thus established the following theorem.

**THEOREM**

*In a four-dimensional space the only tensor density whose components $A^{ij}$ satisfy the conditions $A^{ij} = A^{ji}$, $A^{ij}{}_{|j} = 0$ and $A^{ij}(g_{ab}, g_{ab,c}, g_{ab,cd})$, is given by*

$$A^{ij} = \alpha\sqrt{g}\,(R^{ij} - \tfrac{1}{2}g^{ij}R) + \lambda\sqrt{g}\,g^{ij}$$

*where $\alpha$ and $\lambda$ are constants.*

*Remark 1.* This theorem establishes quite clearly that (3.52) are the only possible second-order Euler-Lagrange expressions in a four-dimensional space obtainable from a scalar density of the form (3.2). Nevertheless, it does *not* establish that $L_0$, defined by (3.49), is the only scalar density of the form (3.2) with this property. In fact, the most general scalar density of the form (3.2) for which the associated Euler-Lagrange equations are of second order has been found for $n \leq 4$ (Lovelock [2]), the relevant scalar density in the four-dimensional case being

$$L = \alpha\sqrt{g}\,R - 2\lambda\sqrt{g} + \beta\varepsilon^{ijkl}R^{ab}{}_{ij}R_{abkl} + \gamma\sqrt{g}\,(R^2 - 4R_i^j R_j^i + R^{ab}{}_{ij}R^{ij}{}_{ab}), \tag{4.36}$$

where $R^{ab}_{\ \ ij} = g^{ah}R_h^{\ b}_{\ ij}$ and $\alpha$, $\lambda$, $\beta$, and $\gamma$ are constants. At first sight one might be tempted to think that the coefficients of $\beta$ and $\gamma$ contribute to the Euler-Lagrange expression. However, it can be shown by direct calculation that

$$\left.\begin{array}{c} E^{rs}(\varepsilon^{ijkl}R^{ab}_{\ \ ij}R_{abkl}) = 0, \\[2mm] E^{rs}(\sqrt{g}(R^2 - 4R_i^j R_j^i + R^{ab}_{\ \ ij}R^{ij}_{\ \ ab})) = 0 \quad \text{for } n = 4. \end{array}\right\} \quad (4.37)$$

For $n > 4$ this problem is, as yet, an open one.

*Remark 2.* The problem pertaining to (4.1)–(4.3) for arbitrary $n$ has been completely settled (Lovelock [3, 5]) the result being

$$A^{lh} = \sqrt{g}\sum_{k=1}^{m-1}\alpha_{(k)}g^{jl}\delta^{hh_1\cdots h_{2k}}_{jj_1\cdots j_{2k}}R^{j_1j_2}_{\ \ h_1h_2}\cdots R^{j_{2k-1}j_{2k}}_{\ \ h_{2k-1}h_{2k}} + \lambda\sqrt{g}\,g^{lh}, \quad (4.38)$$

where $\alpha_{(k)}$, $\lambda$ are arbitrary constants and

$$m = \begin{cases} n/2 & \text{if } n \text{ is even,} \\ (n+1)/2 & \text{if } n \text{ is odd:} \end{cases}$$

Indeed, as in the case of (4.35), this $A^{lh}$ is always the Euler-Lagrange expression corresponding to a suitably chosen scalar density of the form (3.2).

*Remark 3.* Weyl [2] and E. Cartan [1] have shown that, on an $n$-dimensional manifold, the only tensor density satisfying (4.1)–(4.3) *which is linear in $g_{ab,cd}$* is given by (4.35). We can obtain this result immediately from the above analysis in the following way. If $A^{ij}$ is linear in $g_{ab,cd}$ then we must have

$$\frac{\partial^2 A^{ij}}{\partial g_{ab,cd}\,\partial g_{rs,tu}} = 0.$$

However, this is (4.15), but now valid for arbitrary $n$. The analysis which gives rise to (4.35) thus goes through in toto for $n > 2$, and we again recover (4.35). For $n = 2$ one easily sees that $A^{ij;ab,cd} = 0$, which implies that $A^{ij} = \lambda\sqrt{g}\,g^{ij}$. This is consistent with (4.35), since the coefficient of $\alpha$ in (4.35) vanishes identically for $n = 2$ by (7.3.39) and (7.3.40). However, we draw attention to the fact that *in the four-dimensional case the linearity assumption is not required.* In fact, the above analysis can be extended to establish the following result (Lovelock [7]): *In a four-dimensional space the only type (2, 0) tensor density whose components satisfy the conditions*

$$A^{ij} = A^{ij}(g_{ab}, g_{ab,c}, g_{ab,cd}) \quad \text{and} \quad A^{ij}_{\ |j} = 0$$

*is given by*

$$A^{ij} = \alpha\sqrt{g}\left[R^{ij} - \tfrac{1}{2}g^{ij}R\right] + \lambda\sqrt{g}\,g^{ij}.$$

Consequently, not only is the linearity assumption superfluous in the four-dimensional case, but so also is the symmetry assumption (4.3). This result has important applications in the general theory of relativity.

## 8.5    COMBINED VECTOR-METRIC FIELD THEORY

In this section we consider a combination of the vector and metric field theories discussed in Sections 8.2 and 8.3. Here we are concerned with quantities of the type (Rund [3])

$$L(g_{ij}, g_{ij,h}, g_{ij,hk}, \psi_i, \psi_{i,j}) = L_1(g_{ij}, g_{ij,h}, g_{ij,hk}) + L_2(g_{ij}, \psi_i, \psi_{i,j}), \quad (5.1)$$

where $L_1$ and $L_2$ are scalar densities. We adopt the notation

$$\Lambda^{ij} = \frac{\partial L}{\partial g_{ij}}, \qquad \Lambda^{ij,k} = \frac{\partial L}{\partial g_{ij,k}}, \qquad \Lambda^{ij,kl} = \frac{\partial L}{\partial g_{ij,kl}},$$

$$\Psi^i = \frac{\partial L}{\partial \psi_i}, \qquad \Psi^{ij} = \frac{\partial L}{\partial \psi_{i,j}},$$

in which case the *complete* set of Euler-Lagrange equations corresponding to (5.1) is

$$E^{ij}(L) = 0, \quad (5.2)$$

$$E^i(L) = 0, \quad (5.3)$$

where

$$E^{ij}(L) = \frac{d}{dx^l}\left[\Lambda^{ij,l} - \frac{d}{dx^k}(\Lambda^{ij,kl})\right] - \Lambda^{ij},$$

and

$$E^i(L) = \frac{d}{dx^l}\Psi^{il} - \Psi^i.$$

However, because of the special form of (5.1) we have

$$E^{ij}(L) = E^{ij}(L_1) - T^{ij}, \quad (5.4)$$

$$E^i(L) = E^i(L_2), \quad (5.5)$$

where $T^{ij}$ is the energy-momentum tensor density corresponding to $L_2$ introduced in Section 8.2, namely,

$$T^{ij} = \frac{\partial L_2}{\partial g_{ij}}. \quad (5.6)$$

However, (5.4) and (5.5) are not independent. From (5.4) and (3.48), we find

$$E^{ij}(L)_{|j} = -T^{ij}_{|j}. \tag{5.7}$$

By (2.29) and (5.6), we also have

$$T^{ij}_{|j} = \tfrac{1}{2}\{g^{is}\psi_s E^t(L_2)_{|t} + g^{is}F_{js}E^j(L_2)\}. \tag{5.8}$$

From (5.7), (5.8), and (5.5) we thus obtain the identity (Horndeski [2])

$$E^{ij}(L)_{|j} = -\tfrac{1}{2}g^{is}\psi_s E^t(L)_{|t} - \tfrac{1}{2}g^{is}F_{js}E^j(L), \tag{5.9}$$

which clearly displays the relationship between (5.4) and (5.5).

Because of (5.4) and (5.5), the combined Euler-Lagrange equations (5.2) and (5.3) are

$$E^{ij}(L_1) = T^{ij}, \tag{5.10}$$

$$E^i(L_2) = 0. \tag{5.11}$$

As an application of the above analysis, we briefly consider the following example:

$$L = L_0 + \beta\sqrt{g}\, g^{jk}g^{ih}F_{ij}F_{hk} \tag{5.12}$$

where $L_0$ and $F_{ij}$ are defined by (3.49) and (2.17) and $\beta$ is a constant. Clearly (5.12) is of the form (5.1) with

$$L_1 = L_0,$$

$$L_2 = \beta\sqrt{g}\, g^{jk}g^{ih}F_{ij}F_{hk}.$$

From (3.52), (2.33), (2.34), and (5.6) we thus find

$$E^{ij}(L_1) = -\alpha\sqrt{g}\left[\tfrac{1}{2}g^{ij}R - R^{ij}\right] + \lambda\sqrt{g}\,g^{ij},$$

$$E^i(L_2) = 4\beta(\sqrt{g}\,F^{ji})_{|j},$$

and

$$T^{ij} = -2\beta\sqrt{g}\left[F^{ih}F_{kh}g^{kj} - \tfrac{1}{4}g^{ij}(F_{rs}F^{rs})\right], \tag{5.13}$$

so that (5.10) and (5.11) reduce to

$$\alpha[R^{ij} - \tfrac{1}{2}g^{ij}R] + \lambda g^{ij} = -2\beta[F^{ih}F_{kh}g^{kj} - \tfrac{1}{4}g^{ij}(F_{rs}F^{rs})] \tag{5.14}$$

and

$$F^{ij}_{|j} = 0. \tag{5.15}$$

In the four-dimensional case the relations (5.14) and (5.15) are the well-known *Einstein-Maxwell field equations which purport to govern the interaction of gravitational and electromagnetic fields.*

*Remark 1.* It is possible to consider arbitrary scalar densities of the form $L(g_{ij}, g_{ij,h}, g_{ij,hk}, \psi_i, \psi_{i,j})$ without insisting on the particular decomposition (5.1). The subsequent analysis is quite complicated but it can be shown that $E^{ij}(L)$ and $E^i(L)$ are tensor densities of type $(2, 0)$ and $(1, 0)$, respectively, while, moreover, the relation (5.9) is also valid. Furthermore, it should be pointed out that it is possible to obtain (5.14) and (5.15) from a scalar density of this type, namely,

$$L = \alpha\sqrt{g}\,\delta^{abc}_{ijk}\psi_a\psi^i\left[\frac{R^{jk}_{\;\;bc}}{2(\psi_l\psi^l)} + \frac{2\psi^j_{|b}\psi^k_{|c}}{(\psi_i\psi^i)^2}\right] - 2\lambda\sqrt{g} + \beta\sqrt{g}\,F_{ij}F_{hk}g^{jk}g^{ih},$$

where $\psi^j = g^{jk}\psi_k$ [see (7.6.25)].

*Remark 2.* In view of the physical significance of the Einstein-Maxwell field equations, an important problem is to determine the conditions which ensure the inevitability of (5.14) and (5.15). We draw attention to the fact that $L$, given by (5.12), is a scalar density of the general type

$$L = L(g_{ij}, g_{ij,k}, g_{ij,kh}, \psi_{i,j}), \tag{5.16}$$

while (5.14) implies that

$$E^{ij}(L) = E^{ij}(g_{ab}, g_{ab,c}, g_{ab,cd}, \psi_{a,b}). \tag{5.17}$$

With these comments in mind, we cite the following results (Lovelock [12]): If $n = 4$ the only scalar density of the type (5.16) for which (5.17) is valid is

$$L = \alpha\sqrt{g}\,R + \mu\sqrt{g}\,(R^2 - 4R_{ij}R^{ij} + R^{ij}_{\;\;kl}R^{kl}_{\;\;ij}) + \rho\varepsilon^{ijkl}R^{ab}_{\;\;ij}R_{abkl} + \mathscr{L} \tag{5.18}$$

where $\alpha, \mu, \rho$ are constants and $\mathscr{L}$ is a scalar density of the type

$$\mathscr{L} = \mathscr{L}(g_{ab}, \psi_{a,b}). \tag{5.19}$$

The relation (5.18) is clearly of the special form (5.1) and, in this case, the equation (5.10) reduces to

$$\alpha\sqrt{g}\,(R^{ij} - \tfrac{1}{2}g^{ij}R) = \frac{\partial\mathscr{L}}{\partial g_{ij}}. \tag{5.20}$$

In order to restrict $\mathscr{L}$ we can proceed in two different ways. On the one hand (Lovelock [12]) we could demand that

$$\frac{\partial\mathscr{L}}{\partial g_{ij}} = T^{ij} - \lambda g^{ij}, \tag{5.21}$$

where $T^{ij}$ is defined by (5.13), and then attempt to find $\mathscr{L}$. This would ensure that we would obtain (5.14). However, in this case (5.11) may not reduce to

(5.15). On the other hand (Lovelock [10]) we could demand that

$$E^i(\mathscr{L}) = 4\beta(\sqrt{g}\, F^{ji})_{|j}, \tag{5.22}$$

in which case we would obtain (5.15), but (5.20) may not reduce to (5.14). It turns out that (5.21) *and* (5.22) *are equivalent and they each imply that*

$$\mathscr{L} = \beta\sqrt{g}\, F_{ij}F_{hk}g^{jk}g^{ih} + \gamma\varepsilon^{ijkl}F_{ij}F_{kl} - 2\lambda\sqrt{g}, \tag{5.23}$$

where $\gamma$ is a constant, and so in either case we obtain the Einstein-Maxwell equations, the coefficient of $\gamma$ satisfying the Euler-Lagrange equations identically.

## PROBLEMS

8.1   Show that in a $V_4$ for which $ds^2 = -\sum_{\alpha=1}^{3} dx^\alpha\, dx^\alpha + (dx^4)^2$ where $x^4 = ct$ ($c = $ constant), if

$$(F^{ij}) = \begin{pmatrix} 0 & H_3 & -H_2 & E_1 \\ -H_3 & 0 & H_1 & E_2 \\ H_2 & -H_1 & 0 & E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{pmatrix}$$

then Maxwell's equations in vacuo (see Problem 4.5) can be expressed in the form $F^{ij}_{\ |j} = 0$ together with $F_{ij|h} + F_{hi|j} + F_{jh|i} = 0$.

8.2   If $L(g_{ij}, \psi_i, \psi_{i,j}) = \varepsilon^{ijkl}F_{ij}F_{kl}$ in $V_4$, show that $E^i(L)$, defined by (2.21), vanishes identically. By noting that $\varepsilon^{ijkl}F_{ij}F_{kl} = 2(\varepsilon^{ijkl}F_{ij}\psi_l)_{,k}$ explain this. (*Hint:* See Problem 6.18.)

8.3   If $F_{ij} = -F_{ji}$ and $T^i_j = F^{ih}F_{jh} - \frac{1}{4}\delta^i_j(F^{rs}F_{rs})$ show, in a $V_4$ for which $ds^2 = -\sum_{\alpha=1}^{3}(dx^\alpha)^2 + (dx^4)^2$, that $T^i_j = 0$ if and only if $F_{ij} = 0$. Establish that this result is false if

$$ds^2 = -\sum_{\alpha=1}^{2}(dx^\alpha)^2 + \sum_{\beta=3}^{4}(dx^\beta)^2 \quad \text{or} \quad ds^2 = \pm\sum_{\alpha=1}^{4}(dx^\alpha)^2.$$

(This result could be used in relativity to motivate the signature of space-time being $\pm 2$.)

8.4   If $T^i_j$ is defined as in Problem 8.3 show that

$$T^j_{i|j} = F_{ir}F^{jr}_{\ |j} - \tfrac{1}{2}F^{pq}(F_{ip|q} + F_{pq|i} + F_{qi|p}).$$

Hence show that if $F_{ij}$ is defined by (2.17) then

$$T^j_{i|j} = F_{ir}F^{jr}_{\ |j}.$$

*8.5   If $n$ is even and $F_{ij} = -F_{ji}$ show that

$$\varepsilon^{ih_2\cdots h_n}F_{jh_2}F_{h_3 h_4}\cdots F_{h_{n-1}h_n} = \frac{1}{n}\delta^i_j(\varepsilon^{i_1\cdots i_n}F_{i_1 i_2}\cdots F_{i_{n-1}i_n}).$$

If $n$ is odd and $F_{ij} = -F_{ji}$ show that

$$\varepsilon^{ih_2\cdots h_n}F_{jh_2}F_{h_3 h_4}\cdots F_{h_{n-2}h_{n-1}}\psi_{h_n} = \frac{1}{n-1}\delta^i_j\varepsilon^{i_1\cdots i_n}F_{i_1 i_2}\cdots F_{i_{n-2}i_{n-1}}\psi_{i_n}$$

$$-\frac{1}{n-1}\varepsilon^{ii_2\cdots i_n}F_{i_2 i_3}\cdots F_{i_{n-1}i_n}\psi_j.$$

(Lovelock [9]).

8.6   In $V_4$ how many independent components has $F_{ij}F_{kl} - F_{il}F_{kj} + F_{ik}F_{lj}$ if $F_{ij} = -F_{ji}$? Hence show that

$$F_{ij}F_{kl} - F_{il}F_{kj} + F_{ik}F_{lj} = \tfrac{1}{8}\varepsilon_{ijkl}(\varepsilon^{rstu}F_{rs}F_{tu}),$$

and that

$$\det F_{ij} = \tfrac{1}{64}(\varepsilon^{rstu}F_{rs}F_{tu})^2.$$

8.7   If $F_{ij} = -F_{ji}$ determine real numbers $\alpha, \beta$ so that

$$H^{hk}_{\ lj} = F^{hk}F_{lj} + \alpha F^k_{\ l}F^h_{\ j} + \beta F^k_{\ j}F^h_{\ l}$$

has the properties (4.2.6) where $F^k_{\ l} = g^{kj}F_{jl}$.
Hence show that, if $T^i_j$ is defined as in Problem 8.3, then in $V_4$

$$T^i_i = 0,$$
$$T^i_j T^j_k = \tfrac{1}{4}\delta^i_k(T^a_b T^b_a).$$

[This result, when used in conjunction with (5.14) with $\lambda = 0$, immediately gives rise to the so-called algebraic Rainich conditions $R = 0$ and $R^i_j R^j_k = \tfrac{1}{4}\delta^i_k(R^a_b R^b_a)$. (Rainich [1], Misner and Wheeler [1]).]

*8.8   If $L = L(g_{ij}, \psi_{i,j})$ is a scalar density and

$$\frac{\partial L}{\partial g_{ij}}g_{ij} = 0$$

show that, in $V_4$

$$T^i_i = 0$$
$$T^i_j T^j_k = \tfrac{1}{4}\delta^i_k(T^a_b T^b_a)$$

where

$$T^i_j = \frac{\partial L}{\partial g_{ia}}g_{aj}.$$

Give examples of scalar densities with these properties (Lovelock [8]). (This establishes the existence of additional tensors which have the properties necessary to give rise to the algebraic Rainich conditions mentioned in Problem 8.7).

8.9   If $L$ is a scalar density and $L = L(g_{ij}, \phi, \phi_{,i})$ where $\phi$ is a scalar field, show that the invariance identities are

$$\frac{\partial L}{\partial g_{ij}} + \frac{1}{2}\frac{\partial L}{\partial \phi_{,j}}g^{ih}\phi_{,h} = \frac{1}{2}g^{ij}L.$$

Hence show that

$$\left(\frac{\partial L}{\partial g_{ij}}\right)_{|j} = -\frac{1}{2}g^{ih}\phi_{,h}E(L)$$

where

$$E(L) = -\frac{\partial L}{\partial \phi} + \left(\frac{\partial L}{\partial \phi_{,j}}\right)_{,j}$$

If $\partial L/\partial g_{ij} = 0$ establish that $L = 0$.

8.10  Show that

$$\sqrt{g}\,R = \sqrt{g}\,g^{ih}(\gamma_i{}^r{}_j\gamma_r{}^j{}_h - \gamma_i{}^r{}_h\gamma_r{}^j{}_j) + S^j{}_{,j}$$

where $S^j = S^j(g_{ab}, g_{ab,c})$. Explain how this decomposition can be used to account for the fact that $E^{ij}(\sqrt{g}R)$ contains no third and fourth derivatives of $g_{rs}$.

*8.11  Prove that in $V_4$

$$E^{ij}(\varepsilon^{abrs}R_{rscd}R^{cd}{}_{ab}) = 0,$$

$$E^{ij}(\sqrt{g}(R^2 - 4R_{ij}R^{ij} + R^{ab}{}_{ij}R^{ij}{}_{ab})) = 0.$$

8.12  Prove that if $A^{lh}$ is defined by (4.38) then

$$A^{lh} = A^{hl}$$

$$A^{lh}{}_{|l} = 0.$$

*8.13  If

$$L = -2\sqrt{g}\sum_{k=1}^{m-1}\alpha_{(k)}\delta^{h_1\cdots h_{2k}}_{j_1\cdots j_{2k}}R^{j_1j_2}{}_{h_1h_2}\cdots R^{j_{2k-1}j_{2k}}{}_{h_{2k-1}h_{2k}} - 2\lambda\sqrt{g}$$

show that $\Lambda^{ij,kh}{}_{|h} = 0$. Hence show that

$$E^{lh}(L) = A^{lh}$$

where $A^{lh}$ is defined by (4.38).

8.14  Show that properties 1–3 following (4.14) are not independent by showing that 2 is a consequence of 1 and 3.

*8.15  If the components of a type $(0, 1)$ tensor field are denoted by $\phi_i$ and $\phi_i = \phi_i(g_{rs})$ show that $\phi_i = 0$.

*8.16  If the components of a type $(0, 2)$ tensor field are denoted by $\phi_{ij}$ and $\phi_{ij} = \phi_{ij}(g_{rs})$ show that

$$\phi_{ij} = ag_{ij} \quad \text{if } n > 2,$$

and

$$\phi_{ij} = ag_{ij} + b\sqrt{g}\,\varepsilon_{ij} \quad \text{if } n = 2$$

where $a, b$ are constants (Lovelock [2]).

*8.17  If the components of a type $(0, 3)$ tensor field are denoted by $\phi_{ijk}$ and $\phi_{ijk} = \phi_{ijk}(g_{ab})$ show that

$$\phi_{ijk} = 0 \qquad \text{for } n \neq 3$$

and

$$\phi_{ijk} = a\sqrt{g}\,\varepsilon_{ijk} \quad \text{for } n = 3$$

($a$ a constant).

8.18  If $L = L(g_{ab}, g_{ab,c}, \psi_{a,b})$ and $L$ is a scalar density show that

$$\frac{\partial L}{\partial g_{ab,c}} = 0.$$

*8.19  Show that if $T^{ijkl}$ are the components of a type $(4, 0)$ tensor field for which $T^{ijkl}$ is totally symmetric,

$$T^{ijkl} = T^{ijkl}(g_{ab}, g_{ab,c}, g_{ab,cd})$$

$$T^{ijkl}{}_{|i} = 0$$

then

$$T^{ijkl} = a(g^{ij}g^{kl} + g^{il}g^{jk} + g^{ik}g^{jl})$$

where $a$ is a constant (Bedet and Lovelock [1]). (Compare with Problem 7.39.)

*8.20  Construct all scalar densities of the form $L = L(\psi_i, \psi_{i,j})$ (Lovelock [12]).

*8.21  If $n > 2$ show that the only symmetric tensor $B^{ij}$ for which

$$B^{ij} = B^{ij}(g_{ab}, \psi_a, \psi_{a,b}) \quad \text{and} \quad B^{ij}{}_{|j} = \alpha^{ih}F_{h|j}{}^j$$

where $\alpha^{ih}$ is a tensor and $\alpha^{ih} = \alpha^{ih}(g_{ab}, \psi_a, \psi_{a,b})$ is

$$B^{ij} = a\,T^{ij} + b\,g^{ij}$$

where $T^{ij}$ is defined in Problem 8.3. [This result can be used to establish the uniqueness of the electromagnetic energy-momentum tensor, (Lovelock [11]).]

# Appendix

# TENSORS AND FORMS ON DIFFERENTIABLE MANIFOLDS

The approach to the tensor calculus as presented in the main body of the text may, with some justification, be criticized on the grounds that the concept of a tensor is defined entirely in terms of the transformation properties of its components relative to some coordinate system. It is the aim of this Appendix to outline an alternative approach which does not suffer from this defect. To this end the notion of differentiable manifold is defined with some precision, which also serves to augment the superficial description of Section 3.1. This gives rise to a more sophisticated treatment of the tangent spaces and their duals, and tensors are ultimately defined as multilinear functions on appropriate Cartesian products of such spaces. It is then possible to present a rigorous approach to the calculus of exterior forms, which in turn provides an alternative foundation for the theory of curvature. This theory is pursued to a stage at which it blends in naturally with the theory of the main text.

However, the subject matter of this Appendix does not merely furnish a more rigorous basis for the classical tensor calculus. Whereas the analysis of the latter is usually restricted to certain coordinate neighborhoods, the tools to be developed below are indispensable in the investigation of differentiable manifolds "in the large." This is due to the fact that the new techniques diminish the dependence of the analysis on coordinate systems, and thus allow for the abandonment of such restrictions. Although global differential geometry is not treated in the Appendix, the topics discussed therein should render the study of that subject readily accessible. In this regard the reader is

referred, for instance, to the texts by Bishop and Goldberg [1], Brickell and Clark [1], Goldberg [1], Greub, Halperin, and Vanstone [1], and Kobayashi and Nomizu [1].

## A.1 DIFFERENTIABLE MANIFOLDS; TANGENT AND COTANGENT SPACES

A differentiable manifold may be roughly described as a topological space of a certain kind that can be covered by appropriate coordinate neighborhoods. We shall therefore begin with a discussion of topological spaces.

A *topology* on a set $M$ is a class $T$ of subsets of $M$ that satisfies the following conditions: (1) the intersection of every finite class of sets in $T$ is a set in $T$; (2) the union of any class of sets in $T$ is a set in $T$; (3) the empty set $\varnothing$ and $M$ are in $T$. The pair $(M, T)$ is a *topological space* (usually denoted by $M$), and the sets of $T$ are called the *open sets* of this space. The *closed* sets of $M$ are the complements in $M$ of the open sets. The concepts "open" or "closed" are not mutually exclusive: a set can be open, or closed, or both, or neither. The interior $A^\circ$ of a subset $A \subset M$ is the union of all open sets contained in $A$, and from property (2) it follows that $A^\circ$ is open. The closure $\bar{A}$ of $A$ is the intersection of all closed sets that contain $A$. Clearly $\bar{A}$ is closed. A set is open if it coincides with its interior, and closed if it coincides with its closure. The *boundary* of $A$ is the set $\partial A = \bar{A} - A^\circ$.

A topological space is said to be connected if the only sets that are both open and closed are $\varnothing$ and $M$. It is not difficult to show that $M$ is connected if and only if it is *not* the union of two disjoint nonempty open sets. A subset $A$ of $M$ is said to be *dense* in $M$ if $\bar{A} = M$, and $M$ is *separable* if it contains a countable dense subset. A topological space $M$ is called a *Hausdorff space* if, for every pair of distinct points $p \in M$, $q \in M$, there exist open subsets $A$, $B$ such that $p \in A$, $q \in B$, and $A \cap B = \varnothing$. Any open set containing $p$ is called a *neighborhood* of $p$.

A class $\mathscr{A}$ of subsets of the topological space $M$ is a *cover* of $M$ if $M$ coincides with the union of all $A \in \mathscr{A}$. If each set $A \in \mathscr{A}$ is *open*, the class $\mathscr{A}$ is an *open cover* of $M$. A subclass of a cover that is itself a cover is called a *subcover*. A topological space $M$ is said to be *compact* if every open cover has a finite subcover. It may be shown that every closed and bounded subset of $\mathbf{R}^n$ is compact (this is the Heine-Borel theorem), but $\mathbf{R}^n$ itself is not compact.

A cover $\mathscr{B}$ of a topological space $M$ is a *refinement* of a cover $\mathscr{A}$ of $M$ if for every set $B_r \in \mathscr{B}$ there exists at least one set $A_s \in \mathscr{A}$ such that $B_r \subset A_s$. The cover $\mathscr{A}$ is said to be *locally finite* if for each $p \in M$ there exists an open set $W_p$ such that the set $\{ A \in \mathscr{A} : A \cap W_p \neq \varnothing \}$ is finite (that is, the

neighborhood $W_p$ of $p$ has a nonempty intersection with only a finite number of sets of the cover). A Hausdorff topological space $M$ is *paracompact* if every open cover of $M$ has a locally finite refinement.

Let $M, N$ denote a pair of sets. A map (function) $f$ from $M$ into $N$, denoted by $f: M \to N$, is a rule that assigns to each $p \in M$ an element $q \in N$. One writes $q = f(p)$, and $f(M) = \{ f(p) : p \in M \}$. If $f(M) = N$, the map $f$ is said to be *onto*, or *surjective*. If $f(p) = f(q)$ implies that $p = q$, the map is said to be *one-to-one*, or *injective*. If $f$ is both surjective and injective, it has an inverse $f^{-1}: N \to M$, and $f$ is called a *bijection*. For three sets $M, N, P$, with maps $f: M \to N, g: N \to P$, the composition of $g$ and $f$, denoted by $g \circ f$, is the map $M \to P$ which is obtained by following $f$ by $g$ for all $p \in M$ for which $f(p)$ is in the domain of $g$.

When the sets $M, N$ are topological spaces, their topologies can be used to define the notion of continuity. A map $f: M \to N$ is *continuous* if for every open set $G$ in $N$ the set $f^{-1}(G)$ in $M$ is open. A bijection $f: M \to N$ is called a *homeomorphism* if both $f$ and $f^{-1}: N \to M$ are continuous. Under these circumstances the topological spaces $M, N$ are said to be *homeomorphic* (and $N$ is the homeomorphic image of $M$). A *topological property* of a topological space $M$ is a property that is common to all topological spaces that are homeomorphic to $M$. For instance, compactness is a topological property.

Let $M$ be a topological space, and let $U$ be an open neighborhood of a point $p \in M$. If, for some given integer $n$, there exists a homeomorphism $h: U \to \mathbf{R}^n$ onto an open subset $\mathbf{R}^n$, the pair $(U, h)$ is called a *chart of dimension $n$*, with *coordinate neighborhood $U$*. Under these circumstances $h(p)$ is an $n$-tuple of real numbers, and we shall write $h(p) = (u^1, \ldots, u^n)$. An *atlas* on $M$ is a collection of charts $\{ U_\alpha, h_\alpha \}$, where $\alpha \in I$, the latter denoting an index set, such that the sets $U_\alpha$ constitute an open cover of $M$. A *topological manifold* is a separable Hausdorff space that can be covered by an atlas, and we shall now restrict our attention to spaces of this kind.

Let us consider a pair of charts $(U_1, h_1), (U_2, h_2)$ for which the intersection $W = U_1 \cap U_2 \neq \varnothing$ (See Figure 5, page 56). For some point $p \in W$ let us write $h_1(p) = (u^1, \ldots, u^n)$, and $h_2(p) = (\bar{u}^1, \ldots, \bar{u}^n)$. This gives rise to the homeomorphism $h_2 \circ h_1^{-1}: h_1(W) \to h_2(W)$, by which $(u^1, \ldots, u^n) \to (\bar{u}^1, \ldots, \bar{u}^n)$, with inverse $h_1 \circ h_2^{-1}: h_2(W) \to h_1(W)$ by which $(\bar{u}^1, \ldots, \bar{u}^n) \to (u^1, \ldots, u^n)$. These homeomorphisms define functional relationships between the $n$-tuples $(u^1, \ldots, u^n), (\bar{u}^1, \ldots, \bar{u}^n)$, which we shall respectively represent as

$$\bar{u}^j = \bar{u}^j(u^1, \ldots, u^n), \quad u^j = u^j(\bar{u}^1, \ldots, \bar{u}^n), \quad (j = 1, \ldots, n). \quad (1.1)$$

If the functions that occur in (1.1) are of class $C^k$, the charts $(U_1, h_1), (U_2, h_2)$ are said to be $C^k$-*compatible*.

Let $f: M \to \mathbf{R}$ be a function on $M$, and let $(U_\alpha, h_\alpha)$ be a chart of an atlas such that $U_\alpha$ is contained in the domain of $f$. The function $f \circ h_\alpha^{-1}$ is a real-valued function on the open set $h_\alpha(U_\alpha)$ of $\mathbf{R}^n$, and $f$ is said to be of class $C^k$ if $f \circ h_\alpha^{-1}$ is of class $C^k$. This property is independent of the choice of the chart $(U_\alpha, h_\alpha)$ if the latter is $C^k$-compatible with all other admissible charts. This is certainly the case if all overlapping charts of the atlas are $C^k$-compatible; an atlas of this kind is said to be $C^k$-compatible. We shall restrict ourselves to the case $C^\infty$-compatibility; a $C^\infty$-atlas on $M$ is a collection of $C^\infty$-compatible charts whose domains cover $M$. A $C^\infty$-atlas is said to be *maximal* if it contains each chart that is $C^\infty$-compatible with any one of its charts, and such a $C^\infty$-atlas is said to define a *differentiable structure* on $M$.

A $C^\infty$-*differentiable manifold* is a topological manifold $M$ together with a differentiable structure. The dimension of $M$ is the dimension of its charts.

Many examples of differentiable manifolds can be cited, including $\mathbf{R}^n$, finite-dimensional vector spaces, and smooth surfaces in $E^3$. Also, the group $GL(n, \mathbf{R})$ of all nonsingular $n \times n$ matrices with real entries has a differentiable structure by virtue of the homeomorphism from $GL(n, \mathbf{R})$ onto the open subset of $\mathbf{R}^{n^2}$ of points whose $n^2$ coordinates define nonvanishing determinants. If $M, N$ are differentiable manifolds of dimension $n$ and $m$ respectively, with $C^\infty$-atlases $\{U_\alpha, h_\alpha\}, \{V_\beta, k_\beta\}$, the product manifold $M \times N$ may be endowed with a $C^\infty$-atlas $\{U_\alpha \times V_\beta, h_\alpha \times k_\beta\}$, where $h_\alpha \times k_\beta: U_\alpha \times V_\beta \to \mathbf{R}^n \times \mathbf{R}^m$; thus $M \times N$ is a differentiable manifold.

For a given chart $(U, h)$, and some $p \in U$, we have $h(p) = (u^1, \dots, u^n)$. The *coordinate functions* are the real-valued functions on $U$ that are represented by individual entries in the set $\{u^1, \dots, u^n\}$: these are denoted by $x^j$ $(j = 1, \dots, n)$. Accordingly $x^j = u^j \circ h$ is a map $U \to \mathbf{R}$, and $x^j(p) = u^j$. If we denote the coordinate functions associated with a pair of overlapping charts $(U_1, h_1), (U_2, h_2)$ by $x^j$ and $\bar{x}^j$ respectively, the relations (1.1) give rise to the transformation equations (3.1.1) and their inverses (3.1.2). If the Jacobian (3.1.6) is positive on $U_1 \cap U_2$, the charts are said to be *oriented consistently*. The manifold is *orientable* if it has an atlas that consists entirely of charts that are oriented consistently.

The set of all $C^\infty$ functions $U \to \mathbf{R}$ on a neighborhood $U$ of a point $p \in M$ is denoted by $C_p^\infty$. For any $f \in C_p^\infty$ the function $g = f \circ h^{-1}$ is a map $\mathbf{R}^n \to \mathbf{R}$, that is, a real-valued function of the coordinates. Accordingly the partial derivatives of $f$ with respect to $x^j$ are defined by

$$\partial_j f \equiv \frac{\partial f}{\partial x^j} = \frac{\partial g}{\partial u^j} \circ h = \frac{\partial (f \circ h^{-1})}{\partial u^j} \circ h. \tag{1.2}$$

This is supposed to indicate that $\partial_j f(p)$ is the derivative with respect to $u^j$ of the real-valued function $f \circ h^{-1}$ at the point $h(p)$ of $\mathbf{R}^n$. The definition

(1.2) entails the usual laws of partial differentiation: thus, for all $f, g \in C_p^\infty$, and $a, b \in \mathbf{R}$, one has

$$\partial_j(af + bg) = a\partial_j f + b\partial_j g \quad \text{(linearity)}, \tag{1.3}$$

$$\partial_j(fg) = f(\partial_j g) + g(\partial_j f) \quad \text{(rule for derivations)}. \tag{1.4}$$

We shall now introduce the notion of the *tangent space* $T_p(M)$ at a point $p$ of a differentiable manifold $M$. This space is the set of all maps $X: C_p^\infty \to \mathbf{R}$ that satisfy the following conditions for all $f, g \in C_p^\infty$ and $a, b \in \mathbf{R}$:

$$X(af + bg) = a(Xf) + b(Xg) \quad \text{(linearity)}, \tag{1.5}$$

$$X(fg) = (Xf)g(p) + f(p)(Xg) \quad \text{(rule for derivations)}, \tag{1.6}$$

with vector space operations in $T_p(M)$ defined for $X, Y \in T_p(M)$ by

$$(X + Y)f = Xf + Yf, \quad (aX)f = a(Xf). \tag{1.7}$$

Any element $X \in T_p(M)$ is said to be a *tangent vector* of $M$ at $p$. When (1.3), (1.4) are compared with (1.5), (1.6) it is seen that each differential operator $\partial_j$ is a vector at $p$ (when followed by evaluation at $p$). Moreover, for any constant $c \in \mathbf{R}$ one has, in view of (1.7) and (1.6): $X(c) = cX(1) = c[X(1) + X(1)] = 2cX(1)$, so that $X(c) = 0$ identically. The following theorem is fundamental to the entire development.

### THEOREM

*Relative to a given chart $(U, h)$ with coordinate functions $x^1, \dots, x^n$, any vector $X \in T_p(M)$, with $p \in U$, admits the representation*

$$X = X^j \frac{\partial}{\partial x^j} \quad \text{(summation convention)}, \tag{1.8}$$

*in which the coefficients $X^j \in C_p^\infty$ are uniquely defined by*

$$X^j = Xx^j. \tag{1.9}$$

*Proof.* Without loss of generality we may assume that the chart $(U, h)$ is such that $h(p) = (0, \dots, 0)$. Let $f \in C_p^\infty$, and let $q \in U$ be in the domain of $f$, with $h(q) = (u^1, \dots, u^n)$. First, it is asserted that

$$f(q) = f(p) + x^j(q)f_j(q), \tag{1.10}$$

where

$$f_j(p) = \partial_j f(p). \tag{1.11}$$

This is established as follows. The function $g = f \circ h^{-1}$ is defined on an open set of $\mathbf{R}^n$ that contains the points $(u^1, \dots, u^n), (0, \dots, 0)$. Thus

$$g(u^1,\ldots,u^n) - g(0,\ldots,0) = \int_0^1 \frac{d}{dt}\big[g(tu^1,\ldots,tu^n)\big]\,dt = u^j g_j(u^1,\ldots,u^n),$$

$$(1.12)$$

where

$$g_j(u^1,\ldots,u^n) = \int_0^1 \frac{\partial g(tu^1,\ldots,tu^n)}{\partial u^j}\,dt.$$

Differentiation of (1.12) with respect to $u^k$ yields

$$\frac{\partial g(u^1,\ldots,u^n)}{\partial u^k} = g_k(u^1,\ldots,u^n) + u^j \frac{\partial g_j(u^1,\ldots,u^n)}{\partial u^k},$$

so that, in particular,

$$\frac{\partial g(0,\ldots,0)}{\partial u^j} = g_j(0,\ldots,0).$$

$$(1.13)$$

A set of $n$ functions $f_j = g_j \circ h$ is defined on $U$, for which

$$f_j(q) = g_j(u^1,\ldots,u^n),$$

$$(1.14)$$

and, if it is recalled that, by definition of $g$, one has

$$g(u^1,\ldots,u^n) = f(q),\quad g(0,\ldots,0) = f(p),$$

it is seen that the substitution in (1.12) of (1.14), together with $x^j(q) = u^j$, yields (1.10). Also, by virtue of (1.2), (1.13), and (1.14),

$$\partial_j f(p) = \frac{\partial g}{\partial u^j} \circ h(p) = \frac{\partial g(0,\ldots,0)}{\partial u^j} = g_j(0,\ldots,0) = f_j(p),$$

which is (1.11) as asserted.

Second, for any $X \in T_p(M)$ we now evaluate $Xf(q)$ by means of (1.10), regarding the point $p$ as fixed. Since $X$ is a derivation,

$$Xf(q) = \big(Xx^j(q)\big)f_j(q) + x^j(q)X\big(f_j(q)\big).$$

In this relation we put $u^j = x^j(q) = 0$, so that $q$ coincides with $p$. Because of (1.11) it is thus found that

$$Xf(p) = \big(Xx^j(p)\big)f_j(p) = \big(Xx^j(p)\big)\big(\partial_j f(p)\big),$$

which implies the statements (1.8) and (1.9) of the theorem.

The $n$ vectors $\partial_j = \partial/\partial x^j$ are called coordinate vectors; they are linearly independent. For, should there exist a relation $A^j \partial_j = 0$, an application to the coordinate function $x^k$ would yield $0 = A^j(\partial_j x^k) = A^j \delta_j^k = A^k$, $(k = 1,\ldots,n)$, contrary to hypothesis. This gives rise to the

## COROLLARY

*The $n$ coordinate vectors $\partial_j$ constitute a basis of the n-dimensional tangent space $T_p(M)$ at each point $p \in M$.*

The bases referred to in this corollary are called the *coordinate bases* (also *natural* or *canonical* bases) induced by the chart $(U, h)$.

Let us consider the effect of a coordinate transformation such as (1.1) on the representation (1.8) of the vector $X$. On $U_1 \cap U_2$ we have

$$X = X^h \frac{\partial}{\partial x^h} = \overline{X}^j \frac{\partial}{\partial \overline{x}^j},$$

where

$$X^h = Xx^h,\quad \overline{X}^j = X\overline{x}^j.$$

Because of (1.9) the second relation can be expressed as

$$\overline{X}^j = X^h \frac{\partial \overline{x}^j}{\partial x^h},$$

which is nothing other than the transformation law (3.2.21) of a type $(1,0)$ tensor. Thus

$$X = \overline{X}^j \frac{\partial}{\partial \overline{x}^j} = X^h \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial}{\partial \overline{x}^j} = X^h \frac{\partial}{\partial x^h},$$

and, since the values $X^h$ are arbitrary, it follows that

$$\frac{\partial}{\partial x^h} = \frac{\partial \overline{x}^j}{\partial x^h} \frac{\partial}{\partial \overline{x}^j}.$$

This conclusion suggests quite clearly how one should proceed with the construction of type $(0,1)$ vectors. To this end we consider the space $T_p^*(M)$ that is dual to the vector space $T_p(M)$, this space consisting of all linear functions $\omega: T_p(M) \to \mathbf{R}$ on $T_p(M)$. This space, too, can be endowed with a vector space structure in an obvious way. For any $X \in T_p(M)$ we denote by $\omega(X)$ or $\langle \omega, X \rangle$ its image in $\mathbf{R}$ by $\omega$ (both notations will be used). For *any* $f \in C_p^\infty$ we now define a unique element $df \in T_p^*(M)$ by requiring that

$$\langle df, X \rangle = Xf$$

$$(1.15)$$

for *all* $X \in T_p(M)$.

The significance of this definition becomes apparent immediately upon specialization. First, let us identify the vector $X$ in (1.15) with a coordinate vector $\partial_j$, which gives

$$\langle df, \partial_j \rangle = \partial_j f.$$

$$(1.16)$$

Second, let us identify $f$ with the coordinate function $x^j$ in (1.15), at the same time using (1.9):

$$\langle dx^j, X \rangle = Xx^j = X^j.$$

From this it is evident that $dx^j$, as a linear map: $T_p(M) \to \mathbf{R}$, when applied to $X \in T_p(M)$, has for its value the $j^{th}$ component of $X$. (In this sense $dx^j$ "selects" $X^j$.) In particular, since the $j^{th}$ component of $\partial_k$ is $\delta_k^j$, it follows that

$$\langle dx^j, \partial_k \rangle = \delta_k^j. \tag{1.17}$$

This shows that the elements $\{dx^k : k = 1, \ldots, n\}$ of $T_p^*(M)$ constitute a basis of $T_p^*(M)$ that is dual to the basis $\{\partial_j : j = 1, \ldots, n\}$ of $T_p(M)$. Accordingly any $\omega \in T_p^*(M)$ can be expressed as

$$\omega = \omega_j dx^j. \tag{1.18}$$

Consequently the element $df \in T_p^*(M)$ admits the representation

$$df = f_h dx^h. \tag{1.19}$$

Because of (1.17) this yields

$$\langle df, \partial_j \rangle = f_h \langle dx^h, \partial_j \rangle = f_h \delta_j^h = f_j,$$

so that, by (1.16), $f_j = \partial_j f$, and (1.19) assumes the form

$$df = (\partial_j f) dx^j, \tag{1.20}$$

which is consistent with the customary expression for the "differential" of a function.

When (1.20) is applied to the coordinate functions $\bar{x}^j$ that are related to $x^h$ by (3.1.1), it is seen that

$$d\bar{x}^j = (\partial_h \bar{x}^j) dx^h. \tag{1.21}$$

Thus the coordinate representations of the element (1.18) of $T_p^*(M)$ can be expressed as

$$\omega = \bar{\omega}_j d\bar{x}^j = \omega_j (\partial_h \bar{x}^j) dx^h = \omega_h dx^h,$$

which clearly implies the transformation law (3.2.4) for type $(0, 1)$ tensors. Thus the coefficients $\omega_j$ of $\omega$ in (1.18) are nothing other than the components of a covariant vector. Consequently the dual tangent space $T_p^*(M)$ is called the *cotangent space*, and its elements $\omega$ are referred to as *covectors* or *1-forms*.

Thus far our attention has been restricted to single vectors, that is, to individual elements of a given tangent space $T_p(M)$. In practice one is concerned with vector *fields* on $M$ (or on some subset $N$ of $M$): such a

field is given when a unique element is prescribed in each $T_p(M)$ for all points $p \in M$ (or $p \in N$). Relative to a coordinate system $x^j$ the components of individual vectors are defined precisely as before; in general the components $X^j$ of $X \in T_p(M)$ are functions of the coordinates $x^j$ of $p$. A vector field is said to be a $C^\infty$ field if these are $C^\infty$ functions, as will be supposed henceforth. The set theoretic union of all $T_p(M)$, as $p$ ranges over $M$, is called the *tangent bundle* $T(M)$ of $M$; thus each representative of a vector field is in $T(M)$. Similar definitions apply, of course, to covector fields whose representatives are in the *cotangent bundle* $T^*(M)$, the latter being the set theoretic union of all cotangent spaces $T_p^*(M)$, as $p$ ranges over $M$.

Let $X, Y$ be a pair of vector fields on $M$; at each point $p \in M$ their representatives satisfy the conditions (1.5) and (1.6). The same is obviously true for the combination $aX + bY$, with $a, b \in \mathbf{R}$. Consequently the set $\mathscr{X}(M)$ of all vector fields on $M$ is a vector space (of infinite dimension) over the field $\mathbf{R}$. However, for any pair $f, g \in C_p^\infty$, we have by (1.5) and (1.6),

$$\begin{aligned} XY(fg) &= X((Yf)g + f(Yg)) \\ &= (XYf)g + (Yf)(Xg) + (Xf)(Yg) + f(XYg), \end{aligned} \tag{1.22}$$

which shows that $XY$, although defined as a differential operator, is not a derivation and therefore not in $\mathscr{X}(M)$. This is due to the presence of the second and third terms in (1.22), whose sum happens to be symmetric in $X$ and $Y$. Consequently these terms may be eliminated by an interchange of $X$ with $Y$ in (1.22), followed by subtraction. In terms of the notation

$$[X, Y] = XY - YX, \tag{1.23}$$

it is thus found that

$$[X, Y](fg) = ([X, Y]f)g + f([X, Y]g). \tag{1.24}$$

The *Lie bracket* (1.23) of $X, Y$ is therefore a derivation, and, since it also satisfies the condition (5), it follows that *the Lie bracket of a pair of vector fields is a vector field*.

This fact has significant implications. First, we note that, for $a, b \in \mathbf{R}$, and $X, Y, Z \in \mathscr{X}(M)$, the skew-symmetry condition

$$[X, Y] = -[Y, X], \tag{1.25}$$

and the conditions of bilinearity, namely

$$[aX + bY, Z] = a[X, Z] + b[Y, Z], \quad [X, aY + bZ] = a[X, Y] + b[X, Z], \tag{1.26}$$

are valid as direct consequences of the definition (1.23). Second, for any $f \in C_p^\infty$

$$\begin{aligned}
[X,[Y,Z]]f &= X([Y,Z]f) - [Y,Z](Xf) \\
&= X(Y(Zf)) - X(Z(Yf)) - Y(Z(Xf)) \\
&\quad + Z(Y(Xf)).
\end{aligned} \tag{1.27}$$

Two similar relations are obtained by cyclic permutations of $X, Y, Z$. When these are added to (1.27) it is found that all terms on the right-hand side cancel in pairs. Since $f$ is arbitrary, this implies the *Jacobi identity*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0. \tag{1.28}$$

A vector space $V$ over $\mathbf{R}$ on which there is defined a map $V \times V \to V$ which assigns to each pair $X \in V$, $Y \in V$ a unique element $[X,Y] \in V$ is called a real *Lie algebra* whenever this product satisfies the conditions of skew-symmetry, bilinearity, and the Jacobi identity. Thus the above analysis may be summarized in the statement that *the set $\mathscr{X}(M)$ of vector fields on $M$ is a real Lie algebra.*

## A.2  TENSOR ALGEBRA AND EXTERIOR ALGEBRA

From the theory of the preceding section it is evident that vectors and covectors can be defined without any reference to their components in some coordinate system. Instead, they are introduced as elements of the tangent or cotangent spaces of the underlying manifold. We shall now indicate how this procedure can be extended to encompass arbitrary *type $(r,s)$ tensors*. To this end some elementary notions of multilinear algebra are briefly described.

Our considerations will be based on an $n$-dimensional vector space $V$ and its dual $V^*$. By definition of the latter, any element $\omega \in V^*$ is a linear function $\omega : V \to \mathbf{R}$, whose value for a given $X \in V$ is denoted by $\omega(X) = \langle \omega, X \rangle$. Because of linearity one can also regard $X$ as a linear function on $V^*$, with $\langle \omega, X \rangle = X(\omega)$ when evaluated at $\omega$. In this sense $V$ may be identified with the dual $(V^*)^*$ of $V^*$. (This conclusion depends on the assumption that $\dim V$ is finite.)

Let $X, Y$ be vectors in $V$, so that the pair $(X,Y) \in V \times V$ (Cartesian product). A function $f : V \times V \to \mathbf{R}$ is said to be *bilinear* if it is linear in $X$ (for fixed $Y$), and linear in $Y$ (for fixed $X$). Let $\omega, \pi \in V^*$. The *tensor product* of $\omega$ with $\pi$ is a map $\omega \otimes \pi : V \times V \to \mathbf{R}$ with values $\omega \otimes \pi(X,Y) = \omega(X)\pi(Y)$, which is clearly bilinear. More generally, a function $f : V \times \cdots \times V$ ($s$ copies) $\to \mathbf{R}$ is said to be *$s$-linear* if it is linear in each of its arguments $X_1, \ldots, X_s \in V$. For a given set of $s$ elements $\omega^1, \ldots, \omega^s \in V^*$ the tensor product $\omega^1 \otimes \cdots \otimes \omega^s : V \times \cdots \times V \to \mathbf{R}$ is given by

$$\omega^1 \otimes \cdots \otimes \omega^s(X_1, \ldots, X_s) = \omega^1(X_1) \cdots \omega^s(X_s). \tag{2.1}$$

We shall write

$$V^s \times (V^*)^r = V \times \cdots \times V \times V^* \times \cdots \times V^*,$$
$$(s \text{ copies of } V, \ r \text{ copies of } V^*).$$

If $f : V^s \to \mathbf{R}$ is $s$-linear, and $g : (V^*)^r \to \mathbf{R}$ is $r$-linear, their tensor product is the $(s+r)$-linear map $f \otimes g : V^s \times (V^*)^r \to \mathbf{R}$, for which

$$f \otimes g(X_1, \ldots, X_s, \omega^1, \ldots, \omega^r) = f(X_1, \ldots, X_s)g(\omega^1, \ldots, \omega^r). \tag{2.2}$$

The set of all $(s+r)$-linear functions on $V^s \times (V^*)^r$ is denoted by

$$\begin{aligned}
\mathscr{T}_s^r &= (V^*)^s \otimes V^r \\
&= V^* \otimes \cdots \otimes V^* \otimes V \otimes \cdots \otimes V, \ (s \text{ copies of } V^*, \ r \text{ copies of } V).
\end{aligned}$$

A vector space structure is defined on this set in an obvious manner: if $a \in \mathbf{R}$, and $T, S \in \mathscr{T}_s^r$, one defines

$$\begin{aligned}
(aT)(X_1, \ldots, X_s, \omega^1, \ldots, \omega^r) &= aT(X_1, \ldots, X_s, \omega^1, \ldots, \omega^r), \\
(T+S)(X_1, \ldots, X_s, \omega^1, \ldots, \omega^r) &= T(X_1, \ldots, X_s, \omega^1, \ldots, \omega^r) \\
&\quad + S(X_1, \ldots, X_s, \omega^1, \ldots, \omega^r).
\end{aligned}$$

Any element of $\mathscr{T}_s^r$ is a *type $(r,s)$ tensor*; in particular, elements of $\mathscr{T}_0^1 = V$ and $\mathscr{T}_1^0 = V^*$ are respectively type $(1,0)$ and type $(0,1)$ tensors (vectors, covectors). This definition of type $(r,s)$ tensors coincides with that of Chapter 3 when the vector space $V$ is identified with the tangent space $T_p(M)$ at some fixed point $p$ of the underlying manifold $M$.

Let $\{e_j : j = 1, \ldots, n\}$ denote a basis of $V$, with corresponding dual basis $\{\varepsilon^j : j = 1, \ldots, n\}$ in $V^*$:

$$\varepsilon^j(e_h) = \delta_h^j, \qquad e_h(\varepsilon^j) = \delta_h^j. \tag{2.3}$$

Thus any $X \in V$, $\omega \in V^*$ can be represented as

$$X = X^j e_j, \qquad \omega = \omega_j \varepsilon^j, \tag{2.4}$$

and

$$\varepsilon^j(X) = X^j, \qquad e_j(\omega) = \omega_j. \tag{2.5}$$

(The first of (2.5) can be regarded as a defining equation for $\varepsilon^j \in V^*$ for given $\{e_j\}$ in $V$.) Now let $T$ be some element of $\mathscr{T}_1^1 = V^* \otimes V$, that is, $T$ is a bilinear function on $V \times V^*$. Therefore the value of $T$ for the arguments $X, \omega$ as displayed in (2.4) must necessarily be of the form

$$T(X, \omega) = T_j^h X^j \omega_h. \tag{2.6}$$

Because of (2.5) and (2.2) this is equivalent to:

$$T(X, \omega) = T_j^h \varepsilon^j(X) e_h(\omega) = T_j^h \varepsilon^j \otimes e_h(X, \omega).$$

Consequently any $T \in \mathscr{T}_1^1$ can be expressed as

$$T = T_j^h \varepsilon^j \otimes e_h. \tag{2.7}$$

Also, the $n^2$ elements $\{\varepsilon^j \otimes e_h : j, h = 1, \ldots, n\}$ are linearly independent by virtue of the linear independence of $\{\varepsilon^j\}, \{e_h\}$; they therefore constitute a basis of the $n^2$-dimensional space $\mathscr{T}_1^1$. The coefficients $T_j^h$ that appear in (2.7) are the components of the tensor $T$. In the basis $\{\varepsilon^j \otimes e_h\}$ they are given by

$$T_j^h = T(e_j, \varepsilon^h), \tag{2.8}$$

as is immediately evident from (2.7) and (2.3).

This process may be extended directly to the space $\mathscr{T}_s^r$. The elements

$$\left\{ \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_s} \otimes e_{h_1} \otimes \cdots \otimes e_{h_r} : j_\alpha = 1, \ldots, n; h_\beta = 1, \ldots, r; \right.$$
$$\left. \alpha = 1, \ldots, s; \beta = 1, \ldots, r \right\}$$

constitute a basis of the $n^{r+s}$-dimensional space $\mathscr{T}_s^r$. Any $T \in \mathscr{T}_s^r$ can be represented as

$$T = T^{h_1 \cdots h_r}{}_{j_1 \cdots j_s} \varepsilon^{j_1} \otimes \cdots \otimes \varepsilon^{j_s} \otimes e_{h_1} \otimes \cdots \otimes e_{h_r}, \tag{2.9}$$

where

$$T^{h_1 \cdots h_r}{}_{j_1 \cdots j_s} = T(e_{j_1}, \ldots, e_{j_s}, \varepsilon^{h_1}, \ldots, \varepsilon^{h_r}). \tag{2.10}$$

We are now in a position to turn to *exterior* algebra, which is concerned essentially with skew-symmetric tensors. A tensor $T \in \mathscr{T}_2^0$ is said to be symmetric or skew-symmetric according as $T(Y, X) = T(X, Y)$ or $T(Y, X) = -T(X, Y)$, $(X, Y \in V)$. Clearly this is tantamount to symmetry or skew-symmetry in the subscripts of the components $T_{hj}$ of $T$. Accordingly *any skew-symmetric* tensor can be represented as

$$T = T_{hj} \varepsilon^h \otimes \varepsilon^j = T_{hj} \varepsilon^{hj}, \tag{2.11}$$

where

$$\varepsilon^{hj} = \tfrac{1}{2}(\varepsilon^h \otimes \varepsilon^j - \varepsilon^j \otimes \varepsilon^h). \tag{2.12}$$

Since $\varepsilon^{hj} = -\varepsilon^{jh}$, while being otherwise linearly independent, it follows that the elements (2.12), with $h < j$, define a basis of the vector space $\Lambda^2(V)$ of all skew-symmetric type $(0,2)$ tensors. Clearly $\Lambda^2(V)$ is a subspace of $\mathscr{T}_2^0$ and has dimension $\tfrac{1}{2}n(n-1)$.

Let $\omega, \pi \in V^*$. The *exterior* (or *wedge*) *product* of $\omega$ with $\pi$ is defined as

$$\omega \wedge \pi = \tfrac{1}{2}(\omega \otimes \pi - \pi \otimes \omega). \tag{2.13}$$

If $\omega$ is represented by (2.4), while $\pi$ is similarly given by $\pi = \pi_j \varepsilon^j$, we have by (2.13) and (2.12):

$$\omega \wedge \pi = \tfrac{1}{2}\omega_h \pi_j (\varepsilon^h \otimes \varepsilon^j - \varepsilon^j \otimes \varepsilon^h) = \omega_h \pi_j \varepsilon^{hj}, \tag{2.14}$$

which shows that $\omega \wedge \pi \in \Lambda^2(V)$. Moreover, with $\omega = \varepsilon^h, \pi = \varepsilon^j$, the relation (2.13) assumes the special form

$$\varepsilon^h \wedge \varepsilon^j = \tfrac{1}{2}(\varepsilon^h \otimes \varepsilon^j - \varepsilon^j \otimes \varepsilon^h) = \varepsilon^{hj}. \tag{2.15}$$

Consequently the set $\{\varepsilon^h \wedge \varepsilon^j : h < j\}$ is a basis of $\Lambda^2(V)$.

Again, this construction is readily extended to type $(0, s)$ tensors. A tensor $T$ of this type is *totally skew-symmetric* if its value changes sign when the order of any two of its arguments $X_1, \ldots, X_s \in V$ is reversed. Moreover, with any type $(0, s)$ tensor $T$ one can always associate a totally skew-symmetric tensor $AT$ of the same type, the *alternating operator* $A$ being defined by

$$AT(X_1, \ldots, X_s) = \frac{1}{s!} \sum_{(a_1, \ldots, a_s)} (-1)^\kappa T(X_{a_1}, \ldots, X_{a_s}), \tag{2.16}$$

in which the summation extends over all $s!$ permutations of $(1, \ldots, s)$, and $\kappa = 0$ or $1$ according as the permutation $(a_1, \ldots, a_s)$ is even or odd. Let $\Omega, \Pi$ be two totally skew-symmetric tensors of type $(0, s_1)$ and $(0, s_2)$ respectively. The exterior (or wedge) product of $\Omega$ with $\Pi$ is defined as

$$\Omega \wedge \Pi = A(\Omega \otimes \Pi), \tag{2.17}$$

which is a totally skew-symmetric tensor of type $(0, s_1 + s_2)$ [and which is therefore identically zero whenever $s_1 + s_2 > n$]. The following important properties of the exterior product are direct consequences of the definition (2.17):

(i) Associativity: $\Omega \wedge (\Pi \wedge \Lambda) = (\Omega \wedge \Pi) \wedge \Lambda$;
(ii) (Anti)-Commutativity: $\Pi \wedge \Omega = (-1)^{s_1 s_2} \Omega \wedge \Pi$;
(iii) Distributivity: $(\Omega + \Lambda) \wedge \Pi = \Omega \wedge \Pi + \Lambda \wedge \Pi$ (in which it is assumed that $\Lambda$ is also of type $(0, s_1)$).

The set of all totally skew-symmetric tensors of type $(0, s)$—also called $s$-forms—may be endowed with the structure of a vector space, denoted by $\Lambda^s(V)$, with basis elements

$$\varepsilon^{j_1 \cdots j_s} = \{\varepsilon^{j_1} \wedge \cdots \wedge \varepsilon^{j_s} : j_1 < \cdots < j_s\}. \tag{2.18}$$

Clearly $\dim \Lambda^s(V) = \binom{n}{s}$, which is consistent with the notation $\Lambda^1(V) = V^*, \Lambda^0(V) = \mathbf{R}$. The essence of the entire construction is represented by the sequence $\Lambda^0(V), \Lambda^1(V), \Lambda^2(V), \ldots, \Lambda^{n-1}(V), \Lambda^n(V)$ of vector spaces of dimension $1, n, \tfrac{1}{2}n(n-1), \ldots, n, 1$, respectively. The direct sum

$$\Lambda(V) = \Lambda^0 V \oplus \Lambda^1(V) \oplus \cdots \oplus \Lambda^n(V), \tag{2.19}$$

endowed with the exterior product as defined in (2.17) in an associative algebra over **R**. It is called the *exterior algebra* or *Grassmann algebra* over $V$. Clearly

$$\dim \Lambda(V) = 2^n.$$

## A.3  EXTERIOR CALCULUS ON MANIFOLDS

The construction of the previous section is now applied to the theory of differential forms on an $n$-dimensional differentiable manifold $M$, the role of the vector spaces $V, V^*$ being assumed by the tangent and cotangent spaces $T_p(M), T_p^*(M)$ at each point $p$ of $M$. This entails the construction of the spaces $\mathscr{T}_s^0(T_p(M))$ and $\Lambda^s(T_p(M))$ of type $(0, s)$ tensors and $s$-forms respectively. It is convenient (but by no means necessary) to allow the coordinate basis $\{\partial/\partial x^j\}$ of $T_p(M)$ to play the role of the basis $\{e_j\}$ of $V$, which implies that the basis $\{dx^j\}$ of $T_p^*(M)$ assumes the role of the basis $\{\varepsilon^j\}$ of $V^*$. According to (2.18), the set $\{dx^{j_1} \wedge \cdots \wedge dx^{j_s}: j_1 < \cdots < j_s\}$ is a basis of $\Lambda^s(T_p(M))$; thus any $s$-form $\omega$ at $p$ admits the representation

$$\omega = \omega_{j_1 \ldots j_s} dx^{j_1} \wedge \cdots \wedge dx^{j_s}. \tag{3.1}$$

A *field* of $s$-forms is defined on $M$ (or on a subset $N$ of $M$) if a unique $s$-form is assigned to each point $p \in M$ (or $N$). The components $\omega_{j_1 \ldots j_s}$ in (3.1) are then functions of the coordinates of $p$, and if these are $C^\infty$ functions (as will be assumed henceforth) we have a $C^\infty$ field. The set of all $C^\infty$ fields of $s$-forms on $M$ is denoted by $\Lambda^s(M)$, which is endowed with a vector space structure in the usual manner. In particular, the space $\Lambda^0(M)$ is the set of all $C^\infty$ functions (scalar fields, 0-forms) on $M$, while the algebra $\Lambda(M)$ that results from (2.19) is the *algebra of exterior forms* or *Grassmann algebra* on $M$. The space $\Lambda^s(M)$ will frequently be denoted by $\Lambda^s$.

We are now in a position to define exterior derivatives of forms. According to (1.15) one associates with each 0-form $f \in C_p^\infty$ a unique element $df \in T_p^*(M) = \Lambda^1(T_p(M))$. Thus the operator $d$ can be regarded as a map $d: \Lambda^0 \to \Lambda^1$. This immediately raises the question as to the possibility of extending the domain of $d$ to $\Lambda^1, \Lambda^2, \ldots, \Lambda^n$, and we shall now show that this is indeed the case.

To this end we introduce maps $d: \Lambda^s \to \Lambda^{s+1}$ ($0 \leq s \leq n$) that are subject to the following conditions:

(a) If $f \in \Lambda^0$, then $df(X) = Xf$ [which is merely a restatement of (1.15)];

(b) If $\omega \in \Lambda^s, \pi \in \Lambda^s$, then $d(\omega + \pi) = d\omega + d\pi$;

(c) If $\omega \in \Lambda^s, \pi \in \Lambda^t$, then $d(\omega \wedge \pi) = d\omega \wedge \pi + (-1)^s \omega \wedge d\pi$;

(d) For any $\omega \in \Lambda^s$, $d^2\omega = d(d\omega) = 0$.

From (c) and (d) it follows that $d(dx^{j_1} \wedge \cdots \wedge dx^{j_s}) = 0$ identically. Thus, if $d$ is applied to (3.1), the conditions (b) and (c) being observed, it is found that

$$d\omega = d\omega_{j_1 \ldots j_s} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_s}, \tag{3.2}$$

where $d\omega_{j_1 \ldots j_s}$ is specified by (a) in the form (1.20). It follows that the conditions (a)–(d) completely determine the *exterior derivative* $d\omega$ of any $s$-form $\omega$. (The analytical properties of these derivatives are discussed in Chapter 5.)

We shall now turn to the process of *interior multiplication* of forms. This requires the existence of a $C^\infty$ vector field $X$, at least on a neighborhood $U$ of $M$, which then allows one to associate with any $s$-form $\omega$ at $p \in U$ a unique $(s - 1)$-form at $p$, the latter being denoted by $X \lrcorner \omega$. [The notations $i(X)\omega, i_X\omega$ are also used frequently.] The operation $X \lrcorner$ is subject to the following conditions:

(i) For any function (0-form) $f: U \to \mathbf{R}$, $X \lrcorner f = 0$;

(ii) For any 1-form $\omega$, $X \lrcorner \omega = \omega(X) = \langle \omega, X \rangle$;

(iii) For $\omega \in \Lambda^s, \pi \in \Lambda^s$, $X \lrcorner (\omega + \pi) = X \lrcorner \omega + X \lrcorner \pi$;

(iv) For $\omega \in \Lambda^s, \pi \in \Lambda^t$, $X \lrcorner (\omega \wedge \pi) = (X \lrcorner \omega) \wedge \pi + (-1)^s \omega \wedge (X \lrcorner \pi)$.

For example, if $\omega = \omega_j dx^j$, and $X = X^j \partial_j$, it follows from (iii), (iv), (i), and (ii) that

$$X \lrcorner \omega = X \lrcorner (\omega_j dx^j) = (X \lrcorner \omega_j) dx^j + \omega_j (X \lrcorner dx^j) = \omega_j dx^j(X) = \omega_j X^j, \tag{3.3}$$

as required. It is easily verified in the same manner that, for the general $s$-form (3.1) with totally skew-symmetric coefficients,

$$X \lrcorner \omega = s X^j \varepsilon_{jj_2 \ldots j_s} dx^{j_2} \wedge \cdots \wedge dx^{j_s}. \tag{3.4}$$

The following general rules are direct consequences of the conditions (i)–(iv): if $X, Y$ are vector fields on $M$, and $f, g$ are 0-forms,

$$(fX + gY) \lrcorner \omega = f(X \lrcorner \omega) + g(Y \lrcorner \omega), \tag{3.5}$$

$$X \lrcorner Y \lrcorner \omega + Y \lrcorner X \lrcorner \omega = 0. \tag{3.6}$$

Closely associated with the process of interior multiplication is the concept of Lie derivative. Again it must be assumed that a $C^\infty$ vector field $X$ on $U$ is given. The *Lie derivative of an s-form* with respect to $X$ is defined as

$$\mathscr{L}_X\omega = X \lrcorner d\omega + d(X \lrcorner \omega). \tag{3.7}$$

This is evidently also an $s$-form. [The corresponding definition in terms of local coordinates is sketched in Problems 5.17 and 5.18 on page 178.] For the case of a 0-form $f$ the definition (3.7) yields

$$\mathscr{L}_X f = X \lrcorner df + d(X \lrcorner f) = X \lrcorner df = df(X) = Xf. \tag{3.8}$$

Also, from (iv) it is immediately seen that $\mathscr{L}_X$ is a derivation:

$$\mathscr{L}_X(\omega \wedge \pi) = (\mathscr{L}_X\omega) \wedge \pi + \omega \wedge (\mathscr{L}_X\pi). \tag{3.9}$$

If $\omega$ is replaced by $d\omega$ in (3.7) one obtains

$$\mathscr{L}_X(d\omega) = X \lrcorner d^2\omega + d(X \lrcorner d\omega) = d(X \lrcorner d\omega),$$

while exterior differentiation of (3.7) yields

$$d(\mathscr{L}_X\omega) = d(X \lrcorner d\omega) + d^2(X \lrcorner \omega) = d(X \lrcorner d\omega),$$

from which it follows that $\mathscr{L}_X$ and $d$ commute:

$$\mathscr{L}_X d = d\mathscr{L}_X. \tag{3.10}$$

In order to define the Lie derivative of a vector field $Y$ with respect to a given vector field $X$ we impose the additional condition that $\mathscr{L}_X$ be a derivation also when applied to the form $Y \lrcorner \omega$, that is,

$$\mathscr{L}_X(Y \lrcorner \omega) = (\mathscr{L}_X Y) \lrcorner \omega + Y \lrcorner \mathscr{L}_X\omega. \tag{3.11}$$

This determines $\mathscr{L}_X Y$ uniquely. For, if we put $\omega = df$ for some $f \in \Lambda^0$, we have by (3.8)

$$\mathscr{L}_X(Y \lrcorner df) = \mathscr{L}_X(df(Y)) = \mathscr{L}_X(Yf) = XYf,$$

together with

$$(\mathscr{L}_X Y) \lrcorner df = df(\mathscr{L}_X Y) = (\mathscr{L}_X Y)f,$$

while by (3.10) and (3.8)

$$Y \lrcorner (\mathscr{L}_X df) = Y \lrcorner d(\mathscr{L}_X f) = Y \lrcorner d(Xf) = d(Xf)(Y) = YXf.$$

The substitution of these relations in (3.11), with $\omega = df$, then gives $XYf = (\mathscr{L}_X Y)f + YXf$, from which it follows that

$$\mathscr{L}_X Y = [X, Y]. \tag{3.12}$$

As indicated in Section 4.4, the Lie derivative of type $(r, s)$ tensors is well defined and may in fact be constructed for nontensorial geometrical object fields in terms of limiting processes.

We shall now derive a few formulae that are indispensable for any calculation within this framework. First, let $\pi = g\,df$ for some $f, g \in \Lambda^0$. Then by (2.13),

$$d\pi(X, Y) = dg \wedge df(X, Y) = \tfrac{1}{2}(dg \otimes df - df \otimes dg)(X, Y)$$
$$= \tfrac{1}{2}(dg(X)df(Y) - df(X)dg(Y))$$

But, because of (iv), we also have

$$X \lrcorner d\pi = X \lrcorner dg \wedge df = (X \lrcorner dg)df - dg(X \lrcorner df) = dg(X)df - df(X)dg,$$

so that, by (ii),

$$\langle X \lrcorner d\pi, Y \rangle = Y \lrcorner X \lrcorner \pi = dg(X)df(Y) - df(X)dg(Y),$$

and hence $2d\pi(X, Y) = Y \lrcorner X \lrcorner d\pi$. Since any 1-form $\omega$ can be represented as a sum of terms such as $g\,df$, it follows by linearity that

$$2d\omega(X, Y) = Y \lrcorner X \lrcorner d\omega = \langle X \lrcorner d\omega, Y \rangle \tag{3.13}$$

*for all* $\omega \in \Lambda^1$. Second, by (3.7)

$$\langle X \lrcorner d\omega, Y \rangle = \langle \mathscr{L}_X\omega, Y \rangle - \langle d(X \lrcorner \omega), Y \rangle = Y \lrcorner \mathscr{L}_X\omega - Y\omega(X), \tag{3.14}$$

from which we eliminate the second term by means of (3.11):

$$\langle X \lrcorner d\omega, Y \rangle = \mathscr{L}_X(Y \lrcorner \omega) - (\mathscr{L}_X Y) \lrcorner \omega - Y\omega(X)$$
$$= X\omega(Y) - [X, Y] \lrcorner \omega - Y\omega(X),$$

where, in the second step, we have used (3.8) and (3.12). A comparison of this result with (3.13) yields the important formula

$$2d\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y]) \tag{3.15}$$

for the values of the exterior derivative of a 1-form. Similar, but somewhat more complicated formulae for values of $s$-forms with $s > 1$ can be derived. Also, because of (3.14) and (3.13),

$$\langle \mathscr{L}_X\omega, Y \rangle = \langle X \lrcorner d\omega, Y \rangle + Y\omega(X) = 2d\omega(X, Y) + Y\omega(X).$$

This, together with (3.12) and (3.15), implies that

$$\mathscr{L}_X\langle \omega, Y \rangle = \langle \mathscr{L}_X\omega, Y \rangle + \langle \omega, \mathscr{L}_X Y \rangle, \tag{3.16}$$

which displays the nature of $\mathscr{L}_X$ as a derivation once more.

Let $M, N$ denote a pair of differentiable manifolds, and let $\phi: M \to N$ be a $C^\infty$ map of $M$ into $N$. We shall now present a brief analysis of some effects of $\phi$. To this end let us consider some $f \in C_q^\infty$, where $q = \phi(p)$ for a given point $p \in M$. Clearly $F = f \circ \phi \in C_p^\infty$. An induced map $\phi_*: T_p(M) \to T_q(N)$ is now defined as follows: the image $Y = \phi_* X$ of any $X \in T_p(M)$ is supposed to be such that

$$XF = X(f \circ \phi) = Yf = (\phi_* X)f. \tag{3.17}$$

With this is associated a second induced map $\phi^*: T_q^*(N) \to T_p^*(M)$, called

the *pull-back* of $\phi$ in view of the order; for any $\omega \in T_q^*(N)$ the image $\phi^*\omega \in T_p^*(M)$ is defined by the requirement that, for all $X \in T_p(M)$,

$$(\phi^*\omega)(X) = \omega(\phi_* X). \tag{3.18}$$

This construction is readily extended to $s$-forms with $s > 1$: if $\omega \in \Lambda^s(N)$, and $X_1, \ldots, X_s \in T_p(N)$, the pull-back $\phi^*\Omega$ is such that

$$(\phi^*\Omega)(X_1, \ldots, X_s) = \Omega(\phi_* X_1, \ldots, \phi_* X_s). \tag{3.19}$$

In particular, if $f \in \Lambda^0(N)$, then $\phi^*f = f \circ \phi$.

This situation is readily described as follows in terms of local coordinates. If the coordinate functions on $M$ and $N$ are denoted respectively by $\{x^j\}$ and $\{y^\alpha\}$ $(j, h, \ldots = 1, \ldots, n; \alpha, \beta \ldots = 1, \ldots, m)$, the map $\phi$ is represented by a system of $m$ equations $y^\alpha = y^\alpha(x^j)$, and $F(x^j) = f(y^\alpha(x^j))$. Then for $X = X^j(\partial/\partial x^j) \in T_p(M)$, and $Y = \phi_* X = Y^\alpha(\partial/\partial y^\alpha) \in T_q(N)$, the condition (3.17) yields

$$XF = X^j \frac{\partial F(x)}{\partial x^j} = X^j \frac{\partial f(y)}{\partial y^\alpha} \frac{\partial y^\alpha(x)}{\partial x^j} = Y^\alpha \frac{\partial f(y)}{\partial y^\alpha},$$

which implies that the components of $X$ and $Y$ are related by

$$(\phi_* X)^\alpha = Y^\alpha = \frac{\partial y^\alpha(x)}{\partial x^j} X^j. \tag{3.20}$$

Consequently, for $\omega = \omega_\alpha dy^\alpha \in T_q^*(N)$, and $\phi_* \omega = \pi = \pi_j dx^j \in T_p^*(M)$, the condition (3.18) yields

$$\pi(X) = \pi_j X^j = \omega(\phi_* X) = \omega_\alpha Y^\alpha = \omega_\alpha \frac{\partial y^\alpha(x)}{\partial x^j} X^j,$$

so that

$$(\phi^*\omega)_j = \pi_j = \omega_\alpha \frac{\partial y^\alpha(x)}{\partial x^j}. \tag{3.21}$$

The induced maps as defined above possess some very useful properties, two of which are described below. First, if $Y_1 = \phi_*(X_1), Y_2 = \phi_*(X_2)$, then

$$\phi_*[X_1, X_2] = [\phi_* X_1, \phi_* X_2] = [Y_1, Y_2]. \tag{3.22}$$

This is established as follows. Let us replace $X, Y, f$ in (3.17) by $X_1, Y_1, Y_2 f$, respectively, to obtain

$$X_1(Y_2 f \circ \phi) = (Y_1 Y_2 f),$$

while, according to (3.17), with $X = X_2$, $X_2(f \circ \phi) = Y_2 f$, so that

$$X_1 X_2(f \circ \phi) = (Y_1 Y_2 f).$$

A similar relation results from an interchange of $X_1, X_2$, which is subtracted to yield

$$[X_1, X_2](f \circ \phi) = ([Y_1, Y_2] f).$$

This implies (3.22) by virtue of (3.17). Second, for any $s$-form $\omega$ one has

$$d(\phi^*\omega) = \phi^*(d\omega). \tag{3.23}$$

Since any $s$-form admits a representation such as (3.1), and since $\phi^*$ can be regarded as an algebra homomorphism, it is sufficient to verify (3.23) for $\omega = f \in \Lambda^0(N)$, and $\omega = df \in \Lambda^1(N)$. As regards the first case, we merely observe that

$$d(\phi^*f) = d(f \circ \phi) = \phi^*df. \tag{3.24}$$

For the second case we have, on the one hand, $\phi^*(d\omega) = \phi^*(d^2 f) = 0$, while on the other, because of (3.24), $d(\phi^*\omega) = d(\phi^*df) = d^2(\phi^*f) = 0$, which implies the assertion.

## A.4 CONNECTIONS, TORSION, AND CURVATURE

In this section it will be shown how the basic concepts of the classical tensor calculus as developed in Chapter 3 can be introduced and analyzed in a coordinate-free manner. We shall begin with the notion of a connection within the context of the Lie algebra $\mathscr{X}(M)$ of vector fields on our differentiable manifold $M$.

A $C^\infty$ *connection* on $M$ is a linear map $\mathscr{X}(M) \times \mathscr{X}(M) \to \mathscr{X}(M)$, denoted by $\nabla: (X, Y) \to \nabla_X Y$, that is defined for any pair $X, Y \in \mathscr{X}(M)$, and subject to the following conditions:

$$\nabla_{fX+gY} = f\nabla_X + g\nabla_Y, \tag{4.1}$$

$$\nabla_X(fY) = (Xf)Y + f\nabla_X Y, \tag{4.2}$$

where $f, g \in \Lambda^0(M)$. In order to reveal the relationship between the vector fields $\nabla_X Y$ and the vector fields constructed in Chapter 3 in terms of covariant derivatives, we must introduce local basis fields $\{e_j: j = 1, \ldots, n\}$ in the tangent spaces $T_p(M)$ of each point $p$ of a coordinate neighborhood $U$ of $M$, with corresponding dual basis fields $\{\varepsilon^j: j = 1, \ldots, n\}$ for $T_p^*(M)$. Since $\nabla_X Y \in \mathscr{X}(M)$ for all $X, Y \in \mathscr{X}(M)$, $\nabla_X e_j \in \mathscr{X}(M)$, so that

$$\nabla_X e_j = \omega_j{}^h(X) e_h, \tag{4.3}$$

in which the coefficients $\omega_j{}^h(X) \in \mathbf{R}$ are linear in $X$, that is, $\omega_j{}^h \in \Lambda^1(M)$. These $n^2$ 1-forms are called *connection 1-forms*. With $X = X^j e_j, Y = Y^j e_j$, we have, in consequence of (4.2) and (4.3):

$$\nabla_X Y = \nabla_X(Y^j e_j) = (XY^i)e_j + Y^j(\nabla_X e_j) = (XY^j + Y^h \omega_h{}^j(X))e_j, \tag{4.4}$$

in which $\omega_h{}^j(X) = X^k \omega_h{}^j(e_k)$ by linearity, so that

$$\nabla_X Y = \left( XY^j + Y^h X^k \omega_h{}^j(e_k) \right) e_j. \tag{4.5}$$

Evidently the connection $\nabla$ is completely determined by the values $\omega_h{}^j(e_k)$ of the connection 1-forms in a given basis. In particular, whenever the basis $\{\varepsilon^j\}$ is identified with the coordinate basis $\{dx^j\}$ on $U$, we shall write

$$\omega_h{}^j = \Gamma_h{}^j{}_k dx^k, \tag{4.6}$$

with values

$$\omega_h{}^j(\partial_k) = \Gamma_h{}^j{}_m dx^m(\partial_k) = \Gamma_h{}^j{}_m \delta_k^m = \Gamma_h{}^j{}_k. \tag{4.7}$$

In this basis $X = X^k \partial_k$, and (4.5) can be expressed as

$$\nabla_X Y = X^k \left( \partial_k Y^j + Y^h \Gamma_h{}^j{}_k \right) \partial_j = \left( X^k Y^j_{|k} \right) \partial_j, \tag{4.8}$$

where the components of the covariant derivative of the vector field $Y$ are defined as in (3.5.4). The tensorial character of such derivatives is a direct consequence of the fact that $\nabla_X Y$ is defined without reference to local coordinates.

For a given connection $\nabla$ on $M$ we define the *torsion* as a map $T: \mathscr{X}(M) \times \mathscr{X}(M) \to \mathscr{X}(M)$, given by

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y], \tag{4.9}$$

for $X, Y \in \mathscr{X}(M)$. The explicit evaluation of this expression in the arbitrary basis $\{e_j\}$ is readily performed. Since (4.5) can be expressed as

$$\nabla_X Y = \left\{ X\varepsilon^j(Y) + \omega_h{}^j(X)\varepsilon^h(Y) \right\} e_j, \tag{4.10}$$

it follows with the aid of (2.13) that

$$\nabla_X Y - \nabla_Y X = \left\{ X\varepsilon^j(Y) - Y\varepsilon^j(X) + \omega_h{}^j(X)\varepsilon^h(Y) - \omega_h{}^j(Y)\varepsilon^h(X) \right\} e_j$$

$$= \left\{ X\varepsilon^j(Y) - Y\varepsilon^j(X) + 2\left( \omega_h{}^j \wedge \varepsilon^h \right)(X,Y) \right\} e_j.$$

But, according to (3.15),

$$2d\varepsilon^j(X,Y) = X\varepsilon^j(Y) - Y\varepsilon^j(X) - \varepsilon^j([X,Y]),$$

so that

$$\nabla_X Y - \nabla_Y X = 2\left\{ d\varepsilon^j(X,Y) + \left( \omega_h{}^j \wedge \varepsilon^h \right)(X,Y) \right\} e_j + [X,Y].$$

Consequently (4.9) is equivalent to

$$\tfrac{1}{2} T(X,Y) = \Omega^j(X,Y), \tag{4.11}$$

where the *torsion 2-forms* are defined as

$$\Omega^j = d\varepsilon^j + \omega_h{}^j \wedge \varepsilon^h. \tag{4.12}$$

[In the coordinate basis $\{dx^j\}$ this yields in terms of (4.6):

$$\Omega^j = d^2 x^j + \Gamma_h{}^j{}_k dx^k \wedge dx^h = -\tfrac{1}{2}\left( \Gamma_h{}^j{}_k - \Gamma_k{}^j{}_h \right) dx^h \wedge dx^k, \tag{4.13}$$

which is the counterpart of the expression (3.4.18).] It should be observed that, for a torsion-free connection, the relation (4.9) reduces to

$$\nabla_X Y - \nabla_Y X = [X,Y]. \tag{4.14}$$

The connection also gives rise to a set of curvature 2-forms. In order to derive the latter, we begin with the so-called *curvature operator* $K(X,Y)$. The latter is defined for any pair $X, Y \in \mathscr{X}(M)$, and assigns to each $Z \in \mathscr{X}(M)$ a vector field defined by

$$K(X,Y)Z = \left( \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} \right) Z. \tag{4.15}$$

By means of (4.10) and (4.3) it is seen that

$$\nabla_X \nabla_Y Z = \nabla_X \left\{ \left( Y\varepsilon^j(Z) + \omega_h{}^j(Y)\varepsilon^h(Z) \right) e_j \right\}$$

$$= X\left( Y\varepsilon^j(Z) + \omega_h{}^j(Y)\varepsilon^h(Z) \right) e_j$$

$$+ \left( Y\varepsilon^j(Z) + \omega_h{}^j(Y)\varepsilon^h(Z) \right) \omega_j{}^l(X) e_l$$

$$= \left\{ XY\varepsilon^j(Z) + \left( X\omega_h{}^j(Y) \right)\varepsilon^h(Z) + \omega_h{}^j(Y) X\varepsilon^h(Z) \right.$$

$$+ \left. \omega_h{}^j(X) Y\varepsilon^h(Z) + \omega_l{}^j(X)\omega_h{}^l(Y)\varepsilon^h(Z) \right\} e_j.$$

It should be noted that the sum of the third and fourth terms in this expression is symmetric in $X$ and $Y$. With the aid of (2.13) it therefore follows that

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \left\{ (XY - YX)\varepsilon^j(Z) + \left( X\omega_h{}^j(Y) - Y\omega_h{}^j(X) \right)\varepsilon^h(Z) \right.$$

$$+ \left. \left( \omega_l{}^j(X)\omega_h{}^l(Y) - \omega_l{}^j(Y)\omega_h{}^l(X) \right)\varepsilon^h(Z) \right\} e_j$$

$$= \left\{ [X,Y]\varepsilon^j(Z) + \left( X\omega_h{}^j(Y) - Y\omega_h{}^j(X) \right. \right.$$

$$+ \left. \left. 2\omega_l{}^j \wedge \omega_h{}^l(X,Y) \right)\varepsilon^h(Z) \right\} e_j. \tag{4.16}$$

Also, as in (4.10)

$$\nabla_{[X,Y]} Z = \left\{ [X,Y]\varepsilon^j(Z) + \omega_h{}^j([X,Y])\varepsilon^h(Z) \right\} e_j.$$

When this, together with (4.16), is substituted in the definition (4.15), it is found that

$$K(X,Y)Z = \left\{ \left( X\omega_h{}^j(Y) - Y\omega_h{}^j(X) - \omega_h{}^j([X,Y]) \right. \right.$$

$$+ \left. \left. 2\omega_l{}^j \wedge \omega_h{}^l(X,Y) \right)\varepsilon^h(Z) \right\} e_j.$$

Because of (2.15) this reduces to

$$\tfrac{1}{2}K(X,Y)Z = \left\{ \Omega_h{}^j(X,Y)\varepsilon^h(Z) \right\} e_j, \tag{4.17}$$

where the *curvature 2-forms* are defined by

$$\Omega_h{}^j = d\omega_h{}^j + \omega_l{}^j \wedge \omega_h{}^l. \tag{4.18}$$

The relations (4.12) and (4.18) are the *equations of structure*, the latter set being identical with (5.6.3).

We shall conclude with the derivation of some fundamentally important consequences of the equations of structure. The covariant exterior derivatives of individual elements of a set $\{\Pi^j : j = 1, \ldots, n\}$ of $s$-forms are defined by

$$D\Pi^j = d\Pi^j + \omega_h{}^j \wedge \Pi^h. \tag{4.19}$$

When this prescription is applied once more, it is seen that

$$
\begin{aligned}
DD\Pi^j &= d(D\Pi^j) + \omega_h{}^j \wedge (D\Pi^h) \\
&= d\left(d\Pi^j + \omega_h{}^j \wedge \Pi^h\right) + \omega_h{}^j \wedge \left(d\Pi^h + \omega_l{}^h \wedge \Pi^l\right) \\
&= d\omega_h{}^j \wedge \Pi^h - \omega_h{}^j \wedge d\Pi^h + \omega_h{}^j \wedge d\Pi^h + \omega_h{}^j \wedge \omega_l{}^h \wedge \Pi^l.
\end{aligned}
$$

Because of (4.18) is this equivalent to

$$DD\Pi^j = \left(d\omega_h{}^j + \omega_l{}^j \wedge \omega_h{}^l\right) \wedge \Pi^h = \Omega_h{}^j \wedge \Pi^h. \tag{4.20}$$

In particular, if $\Pi^j$ is identified with the basis element $\varepsilon^j$, it follows from (4.12) and the definition (4.19) that

$$D\Omega^j = DD\varepsilon^j = \Omega_h{}^j \wedge \varepsilon^h. \tag{4.21}$$

This is a statement concerning 3-forms: when written out in terms of the basis $\{\varepsilon^j \wedge \varepsilon^h \wedge \varepsilon^k : j < h < k\}$ for such forms, an identity is obtained that reduces to the cyclic identity (3.8.6) when the basis $\{\varepsilon^j\}$ is identified with the coordinate basis $\{dx^j\}$.

The covariant exterior derivatives of the curvature 2-forms are given by

$$D\Omega_h{}^j = d\Omega_h{}^j + \omega_l{}^j \wedge \Omega_h{}^l - \omega_h{}^l \wedge \Omega_l{}^j, \tag{4.22}$$

in which we substitute from (4.18) to obtain

$$
\begin{aligned}
D\Omega_h{}^j &= d\left(d\omega_h{}^j + \omega_l{}^j \wedge \omega_h{}^l\right) + \omega_l{}^j \wedge \left(d\omega_h{}^l + \omega_m{}^l \wedge \omega_h{}^m\right) \\
&\quad - \omega_h{}^l \wedge \left(d\omega_l{}^j + \omega_m{}^j \wedge \omega_l{}^m\right) \\
&= d\omega_l{}^j \wedge \omega_h{}^l - \omega_l{}^j \wedge d\omega_h{}^l + \omega_l{}^j \wedge d\omega_h{}^l + \omega_l{}^j \wedge \omega_m{}^l \wedge \omega_h{}^m \\
&\quad - \omega_h{}^l \wedge d\omega_l{}^j - \omega_m{}^j \wedge \omega_l{}^m \wedge \omega_h{}^l.
\end{aligned}
$$

Since all terms cancel in pairs, this implies that

$$D\Omega_h{}^j = 0 \tag{4.23}$$

identically. Again, when this relation is expressed in terms of a basis of 3-forms an identity is obtained that reduces to the general Bianchi identity (3.8.11) in a coordinate basis.

# BIBLIOGRAPHY

R. Adler, M. Bazin, and M. Schiffer

[1]  *Introduction to General Relativity*, McGraw-Hill, New York, 1965.

N. I. Akhiezer

[1]  *Lectures on the Calculus of Variations*, Blaisdell, New York, 1962.

T. Apostol

[1]  *Mathematical Analysis*, Addison-Wesley, Reading, Mass., 1957.

R. Bach

[1]  "Zur Weylschen Relativitätstheorie und der Weylschen Erweiterung des Krümmungs-tensorbegriffs." *Math. Z.* **9** (1921), 110–135.

D. Barton and J. P. Fitch

[1]  "Applications of algebraic manipulation programs in physics." *Rep. Prog. Phys.* **35** (1972), 235–314.

R. A. Bedet and D. Lovelock

[1]  "Symmetric, divergence-free tensors and the Bel-Robinson tensor." *Utilitas Math.* **3** (1973), 285–295.

L. Bel

[1]  "Sur la radiation gravitationelle." *C. R. Acad. Sci. Paris (A)* **247** (1958), 1094–1096.

T. G. Berry

[1]  "The generalized Gauss-Codazzi equations on a subspace of a Riemannian manifold." *Utilitas Math.* **2** (1972), 141–157.

R. L. Bishop and S. I. Goldberg

[1]  *Tensor Analysis on Manifolds*, Macmillan, New York, 1968.

S. Brehmer and H. Haar

[1]  *Differentialformen und Vektoranalysis*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1973.

F. Brickell and R. S. Clark

[1]  *Differentiable Manifolds*, Van Nostrand-Reinhold, London, 1970.

355

H. W. Brinkmann

[1]  "Riemann spaces conformal to Einstein spaces." *Math. Ann.* **91** (1924), 269–278.

H. A. Buchdahl

[1]  "The Hamiltonian derivatives of a class of fundamental invariants." *Quart. J. Math. Oxford* **19** (1948), 150–159.

C. Carathéodory

[1]  *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, Teubner, Leipzig and Berlin, 1935.
[2]  " Über die Variationsrechung bei mehrfachen Integralen." *Acta Sci. Math.* (*Szeged*) **4** (1929), 193–216.
[3]  "Geometrische Optik," *Ergebn. Math.* **4** (1937), 573–680 (Springer, Berlin).

E. Cartan

[1]  "Sur les équations de la gravitation d'Einstein." *J. Math. Pures Appl.* **1** (1922), 141–203.
[2]  *Leçons sur les invariants intégraux*, Hermann, Paris, 1934.
[3]  *Les systèmes différentiels extérieurs et leurs applications géométriques*, *Actualités scientifiques* **994**, Paris, 1945.
[4]  *Leçons sur la géometrie des espaces de Riemann*, Gauthier-Villars, Paris, 1925 (repr. 1951).

H. Cartan

[1]  *Formes différentielles*, Hermann, Paris, 1967.

S. S. Chern
[1]  "A simple intrinsic proof of the Gauss–Bonnet theorem for closed Riemannian manifolds." *Ann. Math.* **45** (1944), 747–752.

E. T. Davies

[1]  "On the infinitesimal deformations of a space." *Ann. Mat. Pura Appl.* (4) **12** (1933–34), 145–151.

W. R. Davis

[1]  *Classical Fields, Particles, and the Theory of Relativity*, Gordon and Breach, New York, 1970.

J. C. du Plessis

[1]  "Tensorial concomitants and conservation laws." *Tensor* **20** (1969), 347–360.

A. S. Eddington
[1]  *The Mathematical Theory of Relativity*, 2nd ed., Cambridge University Press, Cambridge, 1957.

L. P. Eisenhart

[1]  *Riemannian Geometry*, Princeton University Press, Princeton, 1925.
[2]  *Continuous Groups of Transformations*, Princeton University Press, Princeton, 1933; reprinted Dover, New York, 1961.
[3]  *Non-Riemannian geometry*, American Mathematical Society, New York, 1964.

H. Flanders

[1]  *Differential Forms*, Academic Press, London and New York, 1963.

V. Fock

[1]  *The Theory of Space Time and Gravitation*, Macmillan, New York, 1964.

C. F. Gauss

[1]  *Disquisitiones generales circa superficies curvas*, 1827; *Collected Works*, Vol. 4, 217–258 [translation, Princeton, 1902].

I. M. Gelfand and S. V. Fomin

[1]  *Calculus of Variations*, Prentice-Hall, Englewood Cliffs, 1963.

S. I. Goldberg

[1]  *Curvature and Homology*, Academic Press, New York, 1962.

H. Goldstein

[1]  *Classical Mechanics*, Addison-Wesley, Reading, 1950.

A. E. Green and W. Zerna

[1]  *Theoretical Elasticity*, Oxford University Press, Oxford, 1968.

W. Greub, S. Halperin, and R. Vanstone

[1]  *Connections, Curvature and Cohomology*, Vol. I, Academic Press, New York, 1972.

G. Hadley

[1]  *Linear Algebra*, Addison-Wesley, Reading, Mass., 1961.

B. K. Harrison and F. B. Estabrook

[1]  "Geometric approach to invariance groups and solutions of partial differential systems." *J. Math. Phys.* **12** (1971), 653–666.

S. Hawking and G. Ellis

[1]  *The large scale structure of space-time*, Cambridge University Press, Cambridge, 1973.

D. Hilbert

[1]  "Die Grundlagen der Physik." *Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl.* **1915**, 395–407, **1916**, 53–76; *Math. Ann.* **92** (1924), 1–32.

G. W. Horndeski

[1]  "Dimensionally dependent divergences." *Proc. Cambridge Philos. Soc.* **72** (1972), 77–82.
[2]  *Invariant Variational Principles and Field Theories*, unpublished Ph.D. thesis, University of Waterloo, 1973.

G. W. Horndeski and D. Lovelock

[1]  "Scalar-tensor field theories." *Tensor* **24** (1972), 79–92.

W. Israel

[1]  "Differential Forms in General Relativity." *Comm. Dublin Inst. Adv. Studies Ser. A*, No. 19, 1970.

H. Jeffreys

[1]  *Cartesian Tensors*, Cambridge University Press, Cambridge, 1957.

S. Kobayashi and K. Nomizu

[1]  *Foundations of Differential Geometry*, Vol. I, Wiley-Interscience, New York, 1963.

C. Lanczos

[1] "A remarkable property of the Riemann-Christoffel tensor in four-dimensions." *Ann. Math.* **39** (1938), 842–850.

L. D. Landau and E. M. Lifshitz

[1] *The Classical Theory of Fields*, Addison-Wesley, Reading, Mass., 1959.

T. Levi-Civita

[1] *The Absolute Differential Calculus*, Blackie, London, 1927.

D. Lovelock

[1] "The Lanczos identity and its generalizations." *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **42** (1967), 187–194.
[2] "The uniqueness of the Einstein field equations in a four-dimensional space." *Arch. Rational Mech. Anal.* **33** (1969), 54–70.
[3] "Divergence-free tensorial concomitants." *Aequationes Math.* **4** (1970), 127–138.
[4] "Dimensionally dependent identities." *Proc. Cambridge Philos. Soc.* **68** (1970), 345–350.
[5] "The Einstein tensor and its generalizations." *J. Math. Phys.* **12** (1971), 498–501.
[6] "Intrinsic expressions for curvatures of even order of hypersurfaces in a Euclidean space." *Tensor* **22** (1971), 274–276.
[7] "The four-dimensionality of space and the Einstein tensor." *J. Math. Phys.* **13**, (1972), 874–876.
[8] "The algebraic Rainich conditions." *General Relativity and Gravitation* **4** (1973), 149–159.
[9] Mathematical Aspects of Variational Principles in the General Theory of Relativity, unpublished D.Sc. thesis, University of Natal, 1973.
[10] "The uniqueness of the Einstein–Maxwell field equations." *General Relativity and Gravitation* **5** (1974), 399–408.
[11] "The electromagnetic energy-momentum tensor and its uniqueness." *Int. J. Theoret. Phys.* **10** (1974), 59–65.
[12] "Vector-tensor field theories and the Einstein–Maxwell field equations." *Proc. Roy. Soc. London A* **341** (1974), 285–297.

M. A. McKiernan and H. Richards

[1] "Characterization of some tensor concomitants of the metric tensor and vector fields under restricted groups of transformations." *Ann. Polon. Math.* **23** (1970), 139–149.

C. W. Misner and J. A. Wheeler

[1] "Classical Physics as geometry." *Ann. Phys.* **2** (1957), 525–603.

C. Møller

[1] *The Theory of Relativity*, 2nd ed., Clarendon Press, Oxford, 1972.

E. Nelson

[1] *Tensor Analysis*, Princeton University Press, Princeton, 1967.

A. Nijenhuis

[1] "$X_{n-1}$ forming sets of eigenvectors." *Proc. Kon. Ned. Akad. Amsterdam* **54** (1951), 200–212.

E. Noether

[1] "Invariante Variationsprobleme." *Nachr. Ges. Wiss. Göttingen, Math.-phys. Kl.* **1918**, 235–257.

W. K. H. Panofsky and M. Philips

[1] *Classical Electricity and Magnetism*, Addison-Wesley, Reading, Mass., 1955.

W. Pauli

[1] *Theory of Relativity*, Pergamon, London, 1958.

A. Z. Petrov

[1] *Einstein Spaces*, Pergamon, London, 1969.

G. Y. Rainich

[1] "Electrodynamics in the general relativity theory." *Trans. Am. Math. Soc.* **27** (1925), 106–136.

B. Riemann

[1] *Über die Hypothesen, welche der Geometric zu Grunde liegen*, Habilitationschrift, Göttingen, 1854. [*Collected Works*, Dover, New York, 1953, 273–287.]

W. Rindler

[1] *Special Relativity*, Oliver and Boyd, Edinburgh, 1960.

I. Robinson

[1] Unpublished lecture. King's College, London, 1957.

H. Rund

[1] *The Differential Geometry of Finsler Spaces*, Springer, Berlin-Göttingen-Heidelberg, 1959.
[2] *The Hamilton–Jacobi Theory in the Calculus of Variations*, Van Nostrand, London and New York, 1966; revised and augmented reprint, Krieger, New York, 1973.
[3] "Variational problems involving combined tensor fields." *Abh. math. Sem. Univ. Hamburg* **29** (1966), 243–262.
[4] *Invariant theory of variational problems on subspaces of a Riemannian manifold. Hamburger Math. Einzelschr.* (*New Series*) No. 5, 1971.
[5] "Curvature invariants associated with sets of *n* fundamental forms of hypersurfaces of *n*-dimensional Riemannian manifolds." *Tensor* **22** (1971), 163–173.
[6] " Integral formulae on hypersurfaces of a Riemannian manifold." *Ann. Mat. Pura Appl.* (4) **88** (1971), 99–122.
[7] "Local imbedding of Einstein spaces." *Ann. Mat. Pura Appl.* (4) **93** (1972), 99–110.
[8] "A direct approach to Noether's theorem in the calculus of variations." *Utilitas Math.* **2** (1973), 205–214.
[9] "A divergence theorem for Finsler metrics." To appear in *Monatsh. Math.*, 1975.

H. Rund and D. Lovelock

[1] "Variational principles in the general theory of relativity." *Jber. Deutsch. Math.-Verein.* **74** (1972), 1–65.

J. A. Schouten

[1] *Ricci—Calculus*, 2nd ed., Springer, Berlin-Göttingen-Heidelberg, 1954.

E. Schrödinger

[1] *Space-Time Structure*, Cambridge University Press, Cambridge, 1950.

W. Slebodzinski

[1] *Exterior Forms and their Applications*, Polish Scientific Publishers, Warsaw, 1970.

J. J. Stoker

[1] *Differential Geometry*, Wiley-Interscience, New York, 1969.

J. L. Synge

[1] "On the geometry of dynamics." *Philos. Trans. Roy. Soc. London A* **226** (1926), 31–106.
[2] *Geometrical Optics*, Cambridge University Press, Cambridge, 1937.
[3] *Relativity: The Special Theory*, 2nd ed., North-Holland, Amsterdam, 1964.
[4] *Relativity: The General Theory*, North-Holland, Amsterdam, 1971.

J. L. Synge and A. Schild

[1] *Tensor Calculus*, University of Toronto Press, Toronto, 1949.

G. F. J. Temple

[1] *Cartesian Tensors: An Introduction*, Wiley, New York, 1960.

T. Y. Thomas

[1] *Differential Invariants of Generalized Spaces*, Cambridge University Press, Cambridge, 1934.
[2] "Riemann spaces of class one and their characterization." *Acta Math.* **67** (1936), 169–211.

G. E. Uhlenbeck and G. W. Ford

[1] *Lectures on Statistical Mechanics*, American Mathematical Society, Providence, 1963.

O. Veblen

[1] "Normal coordinates for the geometry of paths." *Proc. Nat. Acad. Sci.* **8** (1922), 192–197.
[2] *Invariants of quadratic differential forms*, Cambridge University Press, Cambridge, 1927.

H. D. Wahlquist and F. B. Estabrook

[1] "Prolongation structures of non-linear evolution equations." *J. Math. Phys.* **16** (1975), 1–7.

H. Weyl

[1] "Reine Infinitesimalgeometrie." *Math. Z* **2** (1918), 384–411.
[2] *Space-Time-Matter*, Methuen, London, 1922.

E. T. Whittaker

[1] *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed., Cambridge University Press, Cambridge, 1937; reprinted Dover, New York, 1944.

K. Yano

[1] *Theory of Lie Derivatives*, North-Holland, Amsterdam, 1955.

# INDEX