UNIVERSAL
LIBRARY

OU_162483

UNIVERSAL
LIBRARY

OSMANIA UNIVERSITY LIBRARY

Call No. 517 | G38L Accession No. 42959

Author Gerretsen, John C. H.

Title Lectures on tensor calculus and differential geometry. 1962

This book should be returned on or before the date last marked below.

LECTURES ON TENSOR CALCULUS AND DIFFERENTIAL GEOMETRY

JOHAN C. H. GERRETSEN

LECTURES
ON
TENSOR CALCULUS
AND
DIFFERENTIAL GEOMETRY

P. NOORDHOFF N.V. GRONINGEN
1962

DEDICATED TO

ENRICO BOMPIANI

ON OCCASION OF HIS 70TH BIRTHDAY

© Copyright 1962 by P. Noordhoff, Groningen, The Netherlands
All rights reserved. This book or parts hereof may not be reproduced
in any form without written permission of the publishers

PREFACE

Geometry, as conceived to-day, is fundamentally the study of spatial structure. But space in the modern sense is only weakly related to the space of ordinary experience. It is an abstract set of unspecified things which may be referred to a system of labels which possess a certain mathematical structure. In differential geometry, where local properties of space are investigated, particular stress is laid on those structures which can be manipulated by the techniques of the infinitesimal calculus.

The increase of knowledge in this part of geometry has been marked by a gradual shifting away from empirical observation. Gauss, for instance, was mainly interested in the local properties of surfaces in ordinary three dimensional space and his investigations were closely related to practical applications. Soon after him, however, mathematicians under the leadership of Riemann were already beginning to see their subject in a new light, for they discovered that many results of earlier times could be generalized for entities which have a close analogy to surfaces, but were not restricted to be sets of points in an ordinary space. A great stimulus arose from the development of a particular technique, the tensor calculus of Ricci, which enabled the geometers to state very complicated relations in a concise form. Thus differential geometry evolved more and more to an extensive formalism, an intricate network of formal relations. This process of abstraction, whereby the formal structure is detached from the concrete subject, has been of great importance. It enabled Einstein to state his famous equations describing the field of gravity in presence of matter.

From the point of view of the pure geometer this approach has many disadvantages. Geometric thinking cannot be cultivated by mere formal computations and, therefore, a more imaginative introduction to the subject-matter remains desirable.

This book is designed to provide its readers with a reasonably self-contained introduction to the methods of differential geometry of general manifolds. But we do not study these manifolds as autonomic spaces. On the contrary, it is always understood that they are embedded in a linear space with a Euclidean metric. Thus the manifolds are endowed with a metric structure induced by that of the environing space.

The most natural instrument for describing properties of a linear space is provided by the vector calculus. We need only a few very elementary operations. For the sake of unity and simplicity the foundations in linear algebra have been laid in conformity with the general tendency of the present time. It is interesting to observe how the more intricate notions of tensor calculus have been reflected in simple operations on vectors. Thus the definition of absolute and covariant differentiation, of the Christoffel three index symbols, and of the Riemannian curvature tensor, can be introduced by means of elementary formulas involving vectors.

From this point of view there is a gradual transition from the simple arithmetic of vectors to the intricate formalism of tensor calculus. Moreover, especially in the study of hypersurfaces, the link between the classical differential geometry of surfaces in ordinary space and their analogues in higher dimensional space can be retained. It is the hope of the author of this book that, when a reader has acquainted himself with the tensor calculus as developed along these lines, he will appreciate the direct introduction of this formalism as needed for the study of general spaces endowed with an arbitrary affine, conformal, or metric structure, without reference to an embedding space.

The first three chapters explain the fundamental notions of linear algebra which are used throughout the subsequent chapters. Chapter 4 has been devoted to a first course in tensor algebra. In chapter 5 the entities have been defined which will be studied by the methods of differential geometry. The sixth chapter deals extensively with the theory of curves. The theory centres on the famous Frenet formulas for which many applications have been given. The main subject of chapter 7 is the theory of geodesic differentiation on arbitrary manifolds with particular emphasis on geodesic lines. Also some other applications, e.g., systems of curves and Ricci's coefficients of rotation, have found a place in this chapter. Chapter 8 treats of the fundamental properties of hypersurfaces which are natural generalizations of classical results of surfaces in ordinary space. The theory dealt with in chapter 9 is mainly the discussion of Riemann's curvature tensor and problems which are intimately related to this tensor. In all these chapters the main stress lies on the algebraic treatment. But differential geometry has a Janus head. Not only linear algebra but also the theory of partial differential equations is an essential part of this field of geometry. The last chapter has been devoted to the theory of integrability and some of its classical applications.

The author of this book did not attempt to produce a complete systematic treatise, but he selected some representative topics. There are many ways of presenting the subject. In the present work the exposition is a result of many years of lecturing experience, not only for the benefit of students in pure mathematics but also for those who appreciate mathematics as an instrument for applied research. Hence it is hoped, that the work will prove to be useful to a rather extensive circle of readers. Much that is new in modern mathematics is perhaps only of transitory importance. The beginner should be warned against overestimating the cultus of abstraction in recent mathematical development, which appears under the name of geometry but has only the name in common with this field of mathematics. In the present work the reader has been carried far enough to appreciate the possibilities of the formalism of tensor calculus, but he should not forget that the ultimate goal of all geometric research is the understanding of geometric objects and that geometric thinking is not the same thing as writing down intricate formulas which are only loosely related to geometric intuition.

Finally it should be noticed that concrete illustrations which make differential geometry in ordinary space so attractive, can hardly be given in higher dimensional space. The main value of differential geometry of general manifolds is the method of approach, culminating in the wonderful formalism of tensor calculus.

Groningen – The Netherlands J. C. H. GERRETSEN

CONTENTS

CHAPTER 8. HYPERSURFACES

CHAPTER 9. THEORY OF CURVATURE OF GENERAL MANIFOLDS

CHAPTER 10. THEORY OF INTEGRABILITY

CHAPTER 1

LINEAR VECTOR SPACES

It is assumed that the reader is familiar with the notion and the most elementary properties of a vector space. In this chapter we shall repeat its definition and discuss some theorems which will be used throughout the subsequent chapters. In many cases the longer proofs are omitted to save space. They are standard in linear algebra.

1.1 – Theorems on vector spaces

1.1.1 – Definition of a linear vector space

The concept of *linear vector space* involves three things:

1) An additive group. Its elements will be denoted by bold-face type.

2) The set of all real numbers.

3) A composition $\boldsymbol{xa}$ of an element of the group and a real number a such that $\boldsymbol{xa}$ is also an element of the group and

$$(\boldsymbol{x}+\boldsymbol{y})a = \boldsymbol{x}a+\boldsymbol{y}a, \tag{1.1–1}$$

$$\boldsymbol{x}(a+b) = \boldsymbol{x}a+\boldsymbol{x}b, \tag{1.1–2}$$

$$(\boldsymbol{x}a)b = \boldsymbol{x}(ab), \tag{1.1–3}$$

$$\boldsymbol{x}1 = \boldsymbol{x}. \tag{1.1–4}$$

The element $\boldsymbol{x}a$ is called the *product* of $\boldsymbol{x}$ and a.

From (1.1–3) we may write $\boldsymbol{x}ab$ instead of $(\boldsymbol{x}a)b$ without fear of confusion.

The elements of a vector space $\mathfrak{X}$ are called *vectors*. The neutral element in the additive group of $\mathfrak{X}$ is called the *zero vector* and is denoted by $\boldsymbol{o}$. It is easy to show that

$$\boldsymbol{o} = \boldsymbol{x}0, \tag{1.1–5}$$

where $\boldsymbol{x}$ is an arbitrary vector and 0 is the number zero.

1.1.2 – Linear subspaces

Any non-empty subset of a vector space $\mathfrak{X}$ is called a *linear subspace,* provided the following condition has been fulfilled:

If $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors of the subset, then all vectors of the type $\boldsymbol{x}a+\boldsymbol{y}b$, where a and b are arbitrary real numbers, also belong to the subset.

It is evident that the whole space $\mathfrak{X}$ is a linear subspace of itself, of course an improper one. By a *proper subspace* is understood a subspace of $\mathfrak{X}$ which does not contain all elements of $\mathfrak{X}$.

The set consisting of the vector $\boldsymbol{o}$ *is a linear subspace of any vector space* $\mathfrak{X}$. For, $\boldsymbol{o}a+\boldsymbol{o}b = \boldsymbol{o}(a+b) = \boldsymbol{o}$.

Let $\mathfrak{S}$ denote any non-empty subset of a vector space $\mathfrak{X}$. The smallest linear subspace $\mathfrak{Y}$ including all vectors of $\mathfrak{S}$ is said to be *spanned* by $\mathfrak{S}$. We also say that $\mathfrak{S}$ is a *generating set* of this smallest subspace $\mathfrak{Y}$.

A linear subspace of $\mathfrak{X}$ spanned by a single vector $\boldsymbol{x} \neq \boldsymbol{o}$ is called a *ray*. It consists of all vectors of the type $\boldsymbol{x}a$, where a runs through the set of all real numbers.

Let $\boldsymbol{y} = \boldsymbol{x}b$, $b \neq 0$, be a vector of the ray. It is easy to prove that the ray generated by $\boldsymbol{y}$ is the same as that generated by $\boldsymbol{x}$, i.e., *a ray can be generated by any of its non-zero vectors.*

The set of non-zero vectors of a ray spanned by a vector $\boldsymbol{x}$ may be decomposed into two disjoint classes, viz.:

1) The class of all vectors $\boldsymbol{x}a$ with $a > 0$.

2) The class of all vectors $\boldsymbol{x}b$ with $b < 0$.

Such a division brought about by one of the generating vectors of the ray is termed an *orientation* of the ray. When we replace $\boldsymbol{x}$ by $-\boldsymbol{x}$ we interchange the classes. We shall say that $\boldsymbol{x}$ and $-\boldsymbol{x}$ determine *opposite orientations* of the ray.

It is clear that *a ray has only two orientations*, for if $\boldsymbol{y} \neq \boldsymbol{o}$ is a vector of the ray then $\boldsymbol{y}$ belongs either to the class $\boldsymbol{x}a$ with $a > 0$ or to the class $\boldsymbol{x}b$ with $b < 0$, and the class $\boldsymbol{y}c$ with $c > 0$ coincides with one of these classes.

1.1.3 – The summation convention

In a vector space $\mathfrak{X}$ we take a finite number of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

A vector $\boldsymbol{a}$ which can be written in the form

$$\boldsymbol{a} = \boldsymbol{x}_1 a^1 + \ldots + \boldsymbol{x}_n a^n, \tag{1.1–6}$$

where $a^1, \ldots, a^n$ are real numbers, is called a *linear combination* of the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. In the expression on the right of (1.1–6) the superscripts do not designate exponents, but simply serve to enumerate the numbers, like the subscripts used to distinguish the vectors one from another. All terms on the right of (1.1–6) are of the type $\boldsymbol{x}_\kappa a^\kappa$, where κ has successively the values $1, \ldots, n$. Instead of (1.1–6) we shall write briefly

$$\boxed{\boldsymbol{a} = \boldsymbol{x}_\kappa a^\kappa,} \tag{1.1–7}$$

agreeing to take the sum of all terms of the type written, when κ *has been replaced* *by the symbols 1,..., n.* This convention is the so-called *summation convention* and will be applied in all cases where the same index appears twice, once as a superscript and once as a subscript. Such an index is called a *dummy index* or a *saturated index* and we are free in selecting the Greek letter denoting it, provided we avoid taking for it an index which appears elsewhere in the expression under consideration. Later on we shall have the opportunity to extend the summation convention to the cases where several indices appear.

1.1.4 – LINEAR INDEPENDENCE AND DEPENDENCE

It is easy to see that if the linear subspace $\mathfrak{Y}$ is spanned by a subset $\mathfrak{S}$ of $\mathfrak{X}$ then every vector of $\mathfrak{Y}$ is a linear combination of certain vectors of $\mathfrak{S}$. A very important case is that in which $\mathfrak{S}$ fulfils the condition of linear independence. A subset $\mathfrak{S}$ of the space $\mathfrak{X}$ is called *linearly independent* if any linear combination (1.1–7) of distinct vectors $\boldsymbol{x}_\kappa$ of $\mathfrak{S}$ is the zero vector only in the case that all coefficients a^κ are zero. This definition implies that *the empty subset is always linearly independent.*

If $\mathfrak{S}$ is linearly independent and generates the vector space $\mathfrak{X}$ then $\mathfrak{S}$ is called a *basis* of $\mathfrak{X}$. In that case every vector of $\mathfrak{X}$ is a linear combination of distinct vectors of $\mathfrak{S}$ and the non-vanishing coefficients in this expression are unique.

If $\mathfrak{X}$ has as basis a finite set $\mathfrak{S}$, then the number of elements of $\mathfrak{S}$ depends only on $\mathfrak{X}$, and is called the *dimension* of $\mathfrak{X}$. This number is the maximal number of elements of $\mathfrak{X}$ constituting an independent set (more briefly stated: the maximal number of independent elements of $\mathfrak{X}$) and any linearly independent set $\mathfrak{S}'$ with the same number of elements as $\mathfrak{S}$ is a basis of $\mathfrak{X}$.

If $\mathfrak{Y}$ is a subspace of $\mathfrak{X}$ then the dimension of $\mathfrak{Y}$ does not exceed the dimension of $\mathfrak{X}$. Equality holds if and only if $\mathfrak{X}$ and $\mathfrak{Y}$ coincide.

The space $\mathfrak{X}$ is called *infinite dimensional* if there is not a maximum of linearly independent vectors in $\mathfrak{X}$.

A set which is not linearly independent is called *linearly dependent.* Thus *any set including the zero vector is linearly dependent.* In particular, the only linearly independent subset of a space consisting solely of the vector $\boldsymbol{o}$ is the empty set. This may be considered the generating set of the subspace consisting of the zero vector only. To this subspace we ascribe the dimension zero.

The dimension of a ray is equal to one. For, a non-zero vector is linearly independent, but any two vectors of a ray constitute a linearly dependent set.

Finally we wish to remark that *a basis of a subspace of a finite dimensional space can be extended to a basis of the whole space.*

The simplest example of an n-dimensional vector space is the set of all

n-tuples of real numbers

$$\boldsymbol{x} = (x^1, \ldots, x^n) \tag{1.1–8}$$

with the following definitions for equality, sum and product:

$$(x^1, \ldots, x^n) = (y^1, \ldots, y^n)$$

if and only if $x^\kappa = y^\kappa$, $\kappa = 1, \ldots, n$,

$$(x^1, \ldots, x^n) + (y^1, \ldots, y^n) = (x^1+y^1, \ldots, x^n+y^n),$$
$$(x^1, \ldots, x^n)a = (x^1 a, \ldots, x^n a).$$

It is easy to see that the vectors

$$\boldsymbol{x}_1 = (1, 0, \ldots, 0),$$
$$\cdots\cdots$$
$$\boldsymbol{x}_n = (0, 0, \ldots, 1)$$

constitute a basis.

1.1.5 – Change of a basis

When a basis $\boldsymbol{x}_\kappa$, $\kappa = 1, \ldots, n$, is assigned in an n-dimensional vector space, then each vector $\boldsymbol{a}$ of the space is uniquely determined by a set of n numbers a^κ, $\kappa = 1, \ldots, n$, such that

$$\boldsymbol{a} = \boldsymbol{x}_\kappa a^\kappa. \tag{1.1–9}$$

The numbers a^κ are called the *components* of the vector $\boldsymbol{a}$ with respect to the given basis.

Next we shall investigate the relation between the components of the same vector with respect to different bases. First we wish to make the following remark. The symbols $1, \ldots, n$ serve to enumerate the elements of the basis under consideration. The arithmetical properties of the numbers represented by these symbols are irrelevant. Thus we might also use any other system of different symbols, e.g., $1', \ldots, n'$ or 1, . . ., n.

Now we introduce a second basis consisting of the vectors $\boldsymbol{x}_{1'}, \ldots, \boldsymbol{x}_{n'}$. The components of the vector (1.1–9) with respect to this basis are $a^{1'}, \ldots, a^{n'}$, such that

$$\boldsymbol{a} = \boldsymbol{x}_{\kappa'} a^{\kappa'}. \quad ✓ \tag{1.1–10}$$

The vectors of the second basis have uniquely determined components with respect to the first basis. We may express this by setting

$$\boxed{\boldsymbol{x}_{\kappa'} = \boldsymbol{x}_\kappa q^\kappa_{\kappa'},} \qquad \kappa' = 1', \ldots, n'. \tag{1.1–11}$$

Conversely, the vectors of the first basis are linear combinations of those of the second basis:

$$x_\kappa = x_{\kappa'} q_\kappa^{\kappa'}. \tag{1.1–12}$$

Now we substitute (1.1–11) into (1.1–12). In order to avoid mistakes we must take different symbols for the free (= non-dummy) indices. We proceed in the following way, writing ρ' instead of κ':

$$x_\lambda = x_{\rho'} q_\lambda^{\rho'} = x_\mu q_{\rho'}^\mu q_\lambda^{\rho'}, \tag{1.1–13}$$

where according to the summation convention in the expression on the right the summation should be carried out with respect to the indices μ and ρ'. On the other hand we also have

$$x_\lambda = x_\mu \delta_\lambda^\mu, \tag{1.1–14}$$

where the *Kronecker symbol* δ_λ^μ denotes the number 1 when λ and μ have the same value and the number 0 when λ and μ are different. Since the components of a vector are unique we may infer from (1.1–13) and (1.1–14) that

$$\boxed{q_\lambda^{\rho'} q_{\rho'}^\mu = \delta_\lambda^\mu.} \tag{1.1–15}$$

Thus we have obtained a relation between the components $q_\kappa^{\kappa'}$, and $q_{\kappa'}^\kappa$ occurring in (1.1–11) and (1.1–12). In a similar way we obtain upon inserting (1.1–12) into (1.1–11) the relation

$$q_{\lambda'}^\rho q_\rho^{\mu'} = \delta_{\lambda'}^{\mu'}, \tag{1.1–16}$$

where $\delta_{\lambda'}^{\mu'}$ is again a Kronecker symbol.

On comparing (1.1–9) and (1.1–10) we get

$$x_{\kappa'} a^{\kappa'} = x_\kappa a^\kappa. \tag{1.1–17}$$

But from (1.1–12) we also have

$$x_\kappa a^\kappa = x_{\kappa'} q_\kappa^{\kappa'} a^\kappa$$

and it follows that

$$\boxed{a^{\kappa'} = q_\kappa^{\kappa'} a^\kappa.} \tag{1.1–18}$$

In a similar way we may find an expression for a^κ in terms of $a^{\kappa'}$:

$$a^\kappa = q_{\kappa'}^\kappa a^{\kappa'}. \tag{1.1–19}$$

This result arises from (1.1–18) by interchanging the primed and the unprimed indices. This is natural, for the two bases play the same role in the vector space $\mathfrak{X}$.

The equations (1.1–18) or (1.1–19) express the *law of transformation* of the components of a vector $\boldsymbol{a}$, when a basis of the space is replaced by another basis.

1.2 - Linear operators

1.2.1 - The concept of linear operator

Let $\mathfrak{X}$ and $\mathfrak{Y}$ denote two vector spaces. A map $\mathsf{A} : \mathfrak{X} \to \mathfrak{Y}$ causes every vector $\boldsymbol{x}$ of $\mathfrak{X}$ to correspond to a uniquely determined vector $\boldsymbol{y}$ of $\mathfrak{Y}$. The vector $\boldsymbol{y}$ is called the *image* of $\boldsymbol{x}$ with respect to the map A and is denoted by

$$\boldsymbol{y} = \mathsf{A}\boldsymbol{x} \tag{1.2–1}$$

The map is said to be a *linear operator* if

$$\mathsf{A}(\boldsymbol{u}a+\boldsymbol{v}b) = (\mathsf{A}\boldsymbol{u})a+(\mathsf{A}\boldsymbol{v})b, \tag{1.2–2}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are arbitrary vectors of $\mathfrak{X}$ and a and b are arbitrary numbers.

It is at once clear that *a linear operator maps a vector space $\mathfrak{X}$ upon a linear subspace of the vector space $\mathfrak{Y}$.*

A trivial example of a linear operator is the correspondence $\boldsymbol{x} \to \boldsymbol{x}$ which leaves every vector of $\mathfrak{X}$ unchanged. In this case $\mathfrak{Y}$ coincides with $\mathfrak{X}$. This operator, called the *identity,* will be denoted by the symbol 1 and we may therefore express the correspondence as

$$\boldsymbol{x} = 1\boldsymbol{x}. \tag{1.2–2}$$

A second example is the following. Let c denote a real number. The correspondence $\boldsymbol{x} \to \boldsymbol{x}c$ is linear, for

$$\begin{aligned}(\boldsymbol{u}a+\boldsymbol{v}b)c &= (\boldsymbol{u}a)c+(\boldsymbol{v}b)c = \boldsymbol{u}(ac)+\boldsymbol{v}(bc)\\ &= \boldsymbol{u}(ca)+\boldsymbol{v}(cb) = (\boldsymbol{u}c)a+(\boldsymbol{v}c)b.\end{aligned}$$

This operator will be denoted by c, i.e.,

$$c\boldsymbol{x} = \boldsymbol{x}c. \tag{1.2–3}$$

This means that a number written at the left-hand side of a vector has the same effect as that written on the right-hand side.

1.2.2 - Operator algebra

Let A and B denote two linear operators which map a vector space $\mathfrak{X}$ into a vector space $\mathfrak{Y}$. By the *sum* $\mathsf{A}+\mathsf{B}$ of these operators we mean a map defined by

$$(\mathsf{A}+\mathsf{B})\boldsymbol{x} = \mathsf{A}\boldsymbol{x}+\mathsf{B}\boldsymbol{x}. \tag{1.2–4}$$

It is easy to show that $\mathsf{A}+\mathsf{B}$ is also a linear operator. The commutativity of the addition is trivial:

$$\mathsf{A}+\mathsf{B} = \mathsf{B}+\mathsf{A}. \tag{1.2–5}$$

The operator $-\mathsf{A}$ is defined by

$$(-\mathsf{A})x = -\mathsf{A}x. \tag{1.2–6}$$

Hence the operator $\mathsf{A}+(-\mathsf{A})$ is the operator which causes every vector x of $\mathfrak{X}$ to correspond to the zero vector o of $\mathfrak{X}$. This is therefore the same operator as that expressed by $0x = o$ and thus we may write

$$\mathsf{A}+(-\mathsf{A}) = 0. \tag{1.2–7}$$

Next we consider a linear map A of $\mathfrak{X}$ into $\mathfrak{Y}$ and a linear map B of $\mathfrak{Y}$ into $\mathfrak{Z}$. The *product* BA of the operators A and B is defined by

$$\mathsf{BA}x = \mathsf{B}(\mathsf{A}x) \tag{1.2–8}$$

and turns out to be a linear map of $\mathfrak{X}$ into $\mathfrak{Z}$. Also in the case that $\mathfrak{X}$, $\mathfrak{Y}$ and $\mathfrak{Z}$ coincide the order of the factors A and B cannot be reversed save in exceptional cases. That is to say, a product of two operators is not always commutative. There are exceptions, however. Thus, for instance,

$$\mathsf{A}a = a\mathsf{A}, \tag{1.2–9}$$

for

$$(\mathsf{A}a)x = \mathsf{A}(ax) = \mathsf{A}(xa) = (\mathsf{A}x)a = a(\mathsf{A}x) = (a\mathsf{A})x.$$

It is also evident that

$$-\mathsf{A} = (-1)\mathsf{A}. \tag{1.2–10}$$

It is always true that the product is associative:

$$\mathsf{C}(\mathsf{BA}) = (\mathsf{CB})\mathsf{A}, \tag{1.2–11}$$

since

$$\mathsf{C}(\mathsf{BA})x = \mathsf{C}(\mathsf{B}(\mathsf{A}x)) = \mathsf{CB}(\mathsf{A}x) = (\mathsf{CB})\mathsf{A}x.$$

A straightforward calculation also shows the validity of the two distributive laws:

$$\mathsf{A}(\mathsf{B}+\mathsf{C}) = \mathsf{AB}+\mathsf{AC}, \qquad (\mathsf{A}+\mathsf{B})\mathsf{C} = \mathsf{AC}+\mathsf{BC}. \tag{1.2–12}$$

In all cases, when we speak of a product BA, it is assumed that this product is defined, that is to say, the operator B operates on all vectors of the linear vector space consisting of the vectors $\mathsf{A}x$.

1.2.3 – Invertable operators

A vector x is called *singular* with respect to a linear operator A when its image is the zero vector in the image space. It is clear that the zero vector in $\mathfrak{X}$ is always singular, for $\mathsf{A}o = \mathsf{A}(o0) = (\mathsf{A}o)0 = o$. In these equations the symbol $\boldsymbol{o}$ denotes the zero vector of $\mathfrak{X}$, except the last symbol which denotes the zero vector of $\mathfrak{Y}$.

The set of all singular vectors of a linear operator A *is a linear subspace.* For, if $\boldsymbol{u}$ and $\boldsymbol{v}$ denote singular vectors of $\mathfrak{X}$ with respect to A, then

$$\mathsf{A}(\boldsymbol{u}a + \boldsymbol{v}b) = (\mathsf{A}\boldsymbol{u})a + (\mathsf{A}\boldsymbol{v})b = \boldsymbol{o}a + \boldsymbol{o}b = \boldsymbol{o} + \boldsymbol{o} = \boldsymbol{o}.$$

The space of all singular vectors of an operator is termed the *kernel* of the operator. When the kernel involves vectors different from zero then the operator is also called *singular*. Hence the kernel of a non-singular operator consists only of the zero vector.

Next we shall make the following assumptions:

1) The operator A is non-singular.

2) The vectors $\mathsf{A}\boldsymbol{x}$ span the whole space $\mathfrak{Y}$.

Since the set of all vectors $\mathsf{A}\boldsymbol{x}$ is a linear vector space, this space must coincide with $\mathfrak{Y}$, i.e., to a given vector $\boldsymbol{y}$ of $\mathfrak{Y}$ we can always find a vector $\boldsymbol{x}$ of $\mathfrak{X}$ such that $\boldsymbol{y} = \mathsf{A}\boldsymbol{x}$. This vector is uniquely determined, for from $\mathsf{A}\boldsymbol{x} = \mathsf{A}\boldsymbol{x}^*$ follows $\mathsf{A}(\boldsymbol{x}-\boldsymbol{x}^*) = \boldsymbol{o}$. Hence, according to the first assumption stated above, $\boldsymbol{x}-\boldsymbol{x}^* = \boldsymbol{o}$, i.e., $\boldsymbol{x} = \boldsymbol{x}^*$.

Consequently, to every vector $\boldsymbol{y}$ of $\mathfrak{Y}$ corresponds a uniquely determined vector $\boldsymbol{x}$ of $\mathfrak{X}$ such that (1.2–1) holds. *This correspondence is again linear.* For, when $\boldsymbol{v}$ and $\boldsymbol{y}$ are vectors of $\mathfrak{Y}$ then we can find vectors $\boldsymbol{u}$ and $\boldsymbol{x}$ in $\mathfrak{X}$ such that $\boldsymbol{v} = \mathsf{A}\boldsymbol{u}$, $\boldsymbol{y} = \mathsf{A}\boldsymbol{x}$. It follows that $\boldsymbol{v}a+\boldsymbol{y}b = (\mathsf{A}\boldsymbol{u})a+(\mathsf{A}\boldsymbol{x})b = \mathsf{A}(\boldsymbol{u}a+\boldsymbol{x}b)$ and thus we see that $\boldsymbol{u}a+\boldsymbol{x}b$ corresponds to $\boldsymbol{v}a+\boldsymbol{y}b$.

The map of $\mathfrak{Y}$ into $\mathfrak{X}$ defined in this way is called the *inverse* of A and is denoted by A^{-1}. An operator satisfying the conditions 1) and 2) stated above is called *invertable*. It is clear that in this case the correspondence between $\mathfrak{X}$ and $\mathfrak{Y}$ is one-to-one.

It is also evident that A^{-1} is again invertable and that

$$(\mathsf{A}^{-1})^{-1} = \mathsf{A}, \tag{1.2–13}$$

while

$$\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}\mathsf{A} = 1. \tag{1.2–14}$$

Finally, let A and B denote two invertable operators. Then

$$(\mathsf{B}\mathsf{A})^{-1} = \mathsf{A}^{-1}\mathsf{B}^{-1}. \tag{1.2–15}$$

The reversal in order of the two factors should be noticed. The proof is easy, for if $\boldsymbol{y} = \mathsf{A}\boldsymbol{x}$, $\boldsymbol{u} = \mathsf{B}\boldsymbol{y}$, then $\boldsymbol{u} = \mathsf{B}\mathsf{A}\,\boldsymbol{x}$ and $\boldsymbol{x} = (\mathsf{B}\mathsf{A})^{-1}\boldsymbol{u}$. But also $\boldsymbol{x} = \mathsf{A}^{-1}\boldsymbol{y}$, $\boldsymbol{y} = \mathsf{B}^{-1}\boldsymbol{u}$, i.e., $\boldsymbol{x} = \mathsf{A}^{-1}\mathsf{B}^{-1}\boldsymbol{u}$.

1.2.4 – Linear operators in a finite dimensional vector space

From now on we confine ourselves to the consideration of an operator A

defined throughout an n-dimensional space $\mathfrak{X}$, which maps $\mathfrak{X}$ into itself. The set of all vectors $\mathsf{A}x$ is a linear subspace of $\mathfrak{X}$ and is therefore also finite dimensional. The dimension of this subspace is called the *rank* of the operator.

First we wish to prove:

In a finite dimensional vector space a non-singular operator A *is always invertable, its rank being the dimension of the space.*

Let x_κ, $\kappa = 1, \ldots, n$, denote a basis of the space. In view of the conditions 1) and 2) stated in the previous section it is sufficient to prove that the vectors $\mathsf{A}x_\kappa$, $\kappa = 1, \ldots, n$ are linearly independent. This may be seen in the following way: The relation

$$(\mathsf{A}x_\kappa)\, a^\kappa = o$$

is equivalent to

$$\mathsf{A}(x_\kappa a^\kappa) = o,$$

whence, since A is non-singular:

$$x_\kappa a^\kappa = o$$

and it follows that $a^\kappa = 0$, $\kappa = 1, \ldots, n$.

Next we have:

If r is the rank of the operator A *and s the dimension of the kernel then*

$$r + s = n, \tag{1.2–16}$$

where n is the dimension of the space.

It is always possible to find a basis x_κ, $\kappa = 1, \ldots, n$, in the vector space whose first s elements are included in the kernel (see the remark at the end of section 1.1.4). If A is non-singular, then $s = 0$, but in this case the theorem is trivial. The statement is also true when all vectors of $\mathfrak{X}$ are singular, for then the rank is zero. Let us suppose that $0 < s < n$. We consider the space spanned by the vectors

$$\mathsf{A}x_{s+1}, \ldots, \mathsf{A}x_n, \tag{1.2–17}$$

Since

$$\mathsf{A}a = \mathsf{A}(x_\kappa a^\kappa) = (\mathsf{A}x_\kappa)\, a^\kappa = (\mathsf{A}x_{s+1})\, a^{s+1} + \ldots + (\mathsf{A}x_n)\, a^n$$

the space spanned by the vectors (1.2–17) is the space of all vectors $\mathsf{A}a$. We wish to prove that (1.2–17) is a basis for this space. But from

$$(\mathsf{A}x_{s+1})\, a^{s+1} + \ldots + (\mathsf{A}x_n)\, a^n = o$$

we deduce that

$$x_{s+1} a^{s+1} + \ldots + x_n a^n = x_1 b^1 + \ldots + x_s b^s$$

and it follows that $a^{s+1} = \ldots = a^n = 0$. The rank of A is therefore $r = n - s$ and thus we have proved the theorem.

As a consequence we have:
An operator is non-singular if and only if its rank is equal to the dimension of the space.

1.2.5 – Matrices

Whenever a basis in the vector space $\mathfrak{X}$ is given we can represent an operator A by means of a square array of numbers, a square *matrix.*

The vectors $\mathsf{A}x_\kappa$, $\kappa = 1, \ldots, n$, are represented as

$$\boxed{\mathsf{A}x_\mu = x_\lambda a^\lambda{}_\mu,} \qquad \mu = 1, \ldots, n. \tag{1.2–18}$$

Conversely, an operator is known when we are given the images of the elements of a basis. Thus we see that an operator A is uniquely characterized by the matrix

$$[a^\lambda{}_\mu] = \begin{bmatrix} a^1{}_1 \; a^1{}_2 \ldots a^1{}_n \\ a^2{}_1 \; a^2{}_2 \ldots a^2{}_n \\ \cdots\cdots\cdots \\ a^n{}_1 \; a^n{}_2 \ldots a^n{}_n \end{bmatrix}. \tag{1.2–19}$$

When A and B are two operators such that

$$\mathsf{A}x_\mu = x_\lambda a^\lambda{}_\mu, \qquad \mathsf{B}x_\mu = x_\lambda b^\lambda{}_\mu, \tag{1.2–20}$$

then evidently

$$(\mathsf{A}+\mathsf{B})x_\mu = x_\lambda(a^\lambda{}_\mu + b^\lambda{}_\mu), \tag{1.2–21}$$

The matrix representing $\mathsf{A}+\mathsf{B}$ with respect to the given basis is called the *sum* of the matrices representing A and B respectively. This is written as

$$\boxed{[a^\lambda{}_\mu] + [b^\lambda{}_\mu] = [a^\lambda{}_\mu + b^\lambda{}_\mu].} \tag{1.2–22}$$

In a similar way it may be seen that the matrix representing $c\mathsf{A}$ is the matrix

$$\boxed{c[a^\lambda{}_\mu] = [ca^\lambda{}_\mu].} \tag{1.2–23}$$

Next we consider the product $\mathsf{C} = \mathsf{BA}$. From (1.2–20) we deduce

$$\mathsf{BA}x_\mu = \mathsf{B}x_\rho a^\rho{}_\mu = (\mathsf{B}x_\rho)a^\rho{}_\mu = x_\lambda b^\lambda{}_\rho a^\rho{}_\mu.$$

Hence C is characterized by the matrix $[c^\lambda{}_\mu]$ with

$$c^\lambda{}_\mu = b^\lambda{}_\rho a^\rho{}_\mu, \qquad \lambda, \mu = 1, \ldots, n, \tag{1.2–24}$$

This matrix is called the *product* of the matrices $[b^\lambda{}_\mu]$ and $[a^\lambda{}_\mu]$. It is written as

$$\boxed{[b^\lambda{}_\mu]\,[a^\lambda{}_\mu] = [b^\lambda{}_\rho\, a^\rho{}_\mu]\,.} \qquad (1.2\text{–}25)$$

The *identity* may be written in the form

$$\mathbf{1}\boldsymbol{x}_\mu = \boldsymbol{x}_\lambda \delta^\lambda_\mu, \qquad (1.2\text{–}26)$$

where δ^λ_μ is the Kronecker symbol introduced in section 1.1.5. Accordingly the operator defined by a real number c is characterized by

$$c\boldsymbol{x}_\mu = \boldsymbol{x}_\lambda c\delta^\lambda_\mu. \qquad (1.2\text{–}27)$$

and it is at once clear that

$$[c\delta^\lambda_\mu]\,[a^\lambda{}_\mu] = [c\delta^\lambda_\rho a^\rho{}_\mu] = [ca^\lambda{}_\mu]$$

is the same matrix as (1.2–23).

If the operator A represented by the matrix $[a^\lambda{}_\mu]$ is invertable we shall represent the inverse operator A^{-1} by the matrix

$$\overset{-1}{[a^\lambda{}_\mu]}$$

From (1.2–24) and (1.2–26) it is at once clear that

$$a^\lambda{}_\rho \overset{-1}{a^\rho{}_\mu} = \overset{-1}{a^\lambda{}_\rho}\, a^\rho{}_\mu = \delta^\lambda_\mu. \qquad (1.2\text{–}28)$$

In elementary algebra it is shown that to a square matrix (1.2–19) corresponds a polynomial

$$\det\,[a^\lambda{}_\mu] = \begin{vmatrix} a^1{}_1 \dots a^1{}_n \\ \cdots\cdots \\ a^n{}_1 \dots a^n{}_n \end{vmatrix} \qquad (1.2\text{–}29)$$

called the *determinant* of the matrix. We recall the product rule:

$$\boxed{\det\,[a^\lambda{}_\mu]\,\det\,[b^\lambda{}_\mu] = \det\,([a^\lambda{}_\mu]\,[b^\lambda{}_\mu])} \qquad (1.2\text{–}30)$$

and the theorem:

A square matrix represents then and only then a singular operator when its determinant has the value zero.

1.2.6 – Law of transformation of a matrix

An operator is defined independently of the choice of a basis. A matrix, however, is intimately connected with a given basis and there arises the problem of finding the relation between two matrices representing the same operator with respect to different bases.

We start from

$$\mathsf{A}x_{\mu'} = x_{\lambda'} a^{\lambda'}{}_{\mu'}. \tag{1.2–31}$$

Taking account of (1.1–12) and (1.1–14) we readily find

$$x_{\lambda'} a^{\lambda'}{}_{\mu'} = \mathsf{A}x_{\mu'} = \mathsf{A}(x_\mu q^\mu_{\mu'})$$
$$= (\mathsf{A}x_\mu) q^\mu_{\mu'} = x_\lambda a^\lambda{}_\mu q^\mu_{\mu'} = x_{\lambda'} q^{\lambda'}_\lambda a^\lambda{}_\mu q^\mu_{\mu'},$$

whence

$$\boxed{a^{\lambda'}{}_{\mu'} = a^\lambda{}_\mu q^{\lambda'}_\lambda q^\mu_{\mu'}.} \quad \checkmark \tag{1.2–32}$$

If A is the identity, we have

$$\delta^{\lambda'}_{\mu'} = \delta^\lambda_\mu q^{\lambda'}_\lambda q^\mu_{\mu'} = q^{\lambda'}_\mu q^\mu_{\mu'} = q^{\lambda'}_\rho q^\rho_{\mu'}, \tag{1.2–33}$$

in accordance with (1.1–16).

Since $\det [\delta^{\lambda'}_{\mu'}] = 1$, we also have

$$\det [q^{\lambda'}_\lambda] \det [q^\lambda_{\lambda'}] = 1 \tag{1.2–34}$$

and it follows that $\det [q^{\lambda'}_\lambda] \neq 0$. As a consequence we find from (1.2–32)

$$\det [a^{\lambda'}{}_{\mu'}] = \det [a^\lambda{}_\mu]. \tag{1.2–35}$$

1.2.7 – Characteristic equation

It may happen that the vector $\mathsf{A}x$ is linearly dependent on x, that is to say, there is a number σ such that

$$\mathsf{A}x = x\sigma. \tag{1.2–36}$$

The zero vector always fulfils the condition (1.2–36). If $x \neq o$ is a vector such that there is a number σ for which the equation (1.2–36) is valid this number is called an *eigenvalue* of the operator A and x an *eigenvector* belonging to the eigenvalue σ.

The set of all vectors satisfying (1.2–36) *for a given number* σ *is a linear subspace of the vector space* $\mathfrak{X}$.

For from $\mathsf{A}u = u\sigma$, $\mathsf{A}v = v\sigma$ it follows that $\mathsf{A}(ua+vb) = (\mathsf{A}u)a+(\mathsf{A}v)b = u\sigma a+v\sigma b = (ua+vb)\sigma$.

If σ is an eigenvalue this space contains at least one vector different from zero. In that case the space is called an *eigenspace* belonging to the eigenvalue σ. In particular, the kernel of A is an eigenspace belonging to $\sigma = 0$ when A is singular.

Let us now assume that the space is finite dimensional. We wish to show that the eigenvalues can be evaluated in an algebraic way. In fact, σ is an eigenvalue if and only if the operator

$$\mathsf{A}-\sigma \tag{1.2–37}$$

is singular. Since this operator is represented by the matrix

$$[a^{\lambda}{}_{\mu}-\sigma\delta^{\lambda}_{\mu}]$$

every eigenvalue is a root of the *characteristic equation*

$$\boxed{\det\,[a^{\lambda}{}_{\mu}-\sigma\delta^{\lambda}_{\mu}] = 0.} \tag{1.2–38}$$

This is an algebraic equation of degree n, where n is the dimension of the space. Since the eigenvalues are defined in an invariant way, that is, independent of the choice of a basis, *it is immaterial which basis is taken for the evaluation of the eigenvalues*.

CHAPTER 2

METRIC VECTOR SPACES

A general linear vector space obtains an additional structure by the introduction of a metric. This enables us to measure the length of a vector and to define the angle between two vectors. We are only concerned with metric spaces of a rather special type which closely resemble Euclidean space.

2.1 – Theorems on metric vector spaces

2.1.1 – INTRODUCTION OF A METRIC

A linear vector space is called a *metric vector space* when to every pair of vectors $\boldsymbol{x}$, $\boldsymbol{y}$ of the space a number $\boldsymbol{xy}$ is assigned, the *inner* or *scalar product* of these vectors, such that the following rules hold:

1) The inner multiplication is commutative:

$$\boldsymbol{xy} = \boldsymbol{yx}. \tag{2.1–1}$$

2) The inner product is associative with respect to the multiplication of a vector by a number:

$$(\boldsymbol{xy})c = \boldsymbol{x}(\boldsymbol{y}c). \tag{2.1–2}$$

3) The inner product is distributive with respect to the addition of vectors:

$$\boldsymbol{x}(\boldsymbol{y}+\boldsymbol{z}) = \boldsymbol{xy}+\boldsymbol{xz}.$$

4) The inner product is positive definite:

$$\boldsymbol{x} \neq \boldsymbol{o}, \text{ implies } \boldsymbol{xx} > 0. \tag{2.1–4}$$

A frequently used practical rule is:

$$(\boldsymbol{x}a)(\boldsymbol{y}b) = \boldsymbol{xy}ab. \tag{2.1–5}$$

For

$$(\boldsymbol{x}a)(\boldsymbol{y}b) = ((\boldsymbol{x}a)\boldsymbol{y})b = (\boldsymbol{y}(\boldsymbol{x}a))b = \boldsymbol{yx}ab = \boldsymbol{xy}ab.$$

Another direct consequence is that *the inner product of two vectors vanishes if at least one of the vectors is zero.* For

$$\boldsymbol{xo} = \boldsymbol{x}(\boldsymbol{o}0) = (\boldsymbol{xo})0 = 0.$$

The converse, however, is not true. It may happen that the inner product of two vectors vanishes without at least one of the vectors being zero. It is said that two vectors are *orthogonal* or *perpendicular* when their inner product vanishes. According to the above statement every vector is orthogonal to the zero vector.

Conversely, *when* $\boldsymbol{x}$ *is orthogonal to all vectors of the space then it is the zero vector,* for then $\boldsymbol{xx} = 0$ and this implies $\boldsymbol{x} = \boldsymbol{o}$.

By the *norm* of a vector $\boldsymbol{x}$ is understood the inner product $\boldsymbol{xx}$ of the vector with itself. It is always positive, except when $\boldsymbol{x} = \boldsymbol{o}$, the norm then being 0. The square root of the norm is the *length* or the *absolute value* of the vector and is denoted by

$$|\boldsymbol{x}| = \sqrt{\boldsymbol{xx}}. \tag{2.1–6}$$

A *unit vector* is a vector whose length is unity. Every vector $\boldsymbol{x} \neq \boldsymbol{o}$ can be transformed into a unit vector when it is multiplied by the number $|\boldsymbol{x}|^{-1}$ or $-|\boldsymbol{x}|^{-1}$. This process is called the *normalizing* of the vector.

2.1.2 – The inequality of Schwarz

The following statement is of fundamental importance in the theory of metric spaces:

If $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are arbitrary vectors then*

$$\boxed{|\boldsymbol{xy}| \leqq |\boldsymbol{x}|\,|\boldsymbol{y}|,} \tag{2.1–7}$$

the equality being valid if and only if the vectors are linearly dependent.

This is *Schwarz's inequality.* The statement is trivial when $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{o}$. Without loss of generality we may suppose that $\boldsymbol{x} \neq \boldsymbol{o}$. First we observe that we can find a number ξ such that

$$\boldsymbol{x}(\boldsymbol{x}\xi + \boldsymbol{y}) = 0 \tag{2.1–8}$$

or

$$\boldsymbol{xx}\xi + \boldsymbol{xy} = 0.$$

The number ξ satisfying this equation is evidently

$$\xi = -(\boldsymbol{xy})(\boldsymbol{xx})^{-1}.$$

From (2.1–8) we conclude that

$$0 \leqq (\boldsymbol{x}\xi + \boldsymbol{y})(\boldsymbol{x}\xi + \boldsymbol{y}) = \boldsymbol{y}(\boldsymbol{x}\xi + \boldsymbol{y}) = (\boldsymbol{xy})\xi + \boldsymbol{yy}.$$

On multiplying by $\boldsymbol{xx}$ we readily find

$$0 \leqq -(\boldsymbol{xy})^2 + (\boldsymbol{xx})(\boldsymbol{yy}).$$

Hence

$$(\boldsymbol{xy})^2 \leqq (\boldsymbol{xx})(\boldsymbol{yy}) = |\boldsymbol{x}|^2 |\boldsymbol{y}|^2$$

and this implies (2.1–7). The equality holds if and only if $\boldsymbol{x}\xi + \boldsymbol{y} = \boldsymbol{o}$. This means, however, that $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly dependent. The result can also be put into the form

$$\begin{vmatrix} \boldsymbol{xx} & \boldsymbol{xy} \\ \boldsymbol{xy} & \boldsymbol{yy} \end{vmatrix} \geqq 0. \tag{2.1–9}$$

Schwarz's inequality enables us to define the *angle* θ between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. With reference to the well-known interpretation of the scalar product of vectors in elementary euclidean geometry we define the *angle* between two non-zero vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ as the number θ which is uniquely determined by

$$\boxed{\boldsymbol{xy} = |\boldsymbol{x}|\,|\boldsymbol{y}| \cos\theta,} \qquad 0 \leqq \theta \leqq \pi. \tag{2.1–10}$$

This angle remains unchanged when we multiply both vectors by a positive or by a negative number. But if one of the vectors is multiplied by a positive number and the other by a negative one then the angle θ must be replaced by $\pi - \theta$.

By the angle between two rays we understand the angle between generating vectors. The angle θ changes into $\pi - \theta$ when the orientation of just one ray is changed. That means that only the angle between orientated rays is uniquely determined.

Sometimes it is convenient to define the angle between two vectors when at least one of them is zero. In that case we shall say that the angle is $\frac{1}{2}\pi$ in accordance with the fact that the angle between two orthogonal vectors is always $\frac{1}{2}\pi$.

Finally we wish to evaluate the determinant (2.1–9) in terms of the angle θ. From (2.1–10) we have

$$(\boldsymbol{xx})(\boldsymbol{yy}) - (\boldsymbol{xy})^2 = (\boldsymbol{xx})(\boldsymbol{yy}) - (\boldsymbol{xx})(\boldsymbol{yy}) \cos^2\theta.$$

Hence

$$\begin{vmatrix} \boldsymbol{xx} & \boldsymbol{xy} \\ \boldsymbol{xy} & \boldsymbol{yy} \end{vmatrix} = |\boldsymbol{x}|^2 |\boldsymbol{y}|^2 \sin^2\theta. \tag{2.1–11}$$

2.1.3 – Gram's determinant

A determinant of the type (2.1–9) can easily be generalized to the case of an arbitrary number of vectors and, as we shall see presently, provides a useful test for linear dependence or independence.

Let $\boldsymbol{x}_\kappa$, $\kappa = 1, \ldots, n$, be a set of n vectors. The determinant

$$\begin{vmatrix} \boldsymbol{x}_1\boldsymbol{x}_1 & \ldots & \boldsymbol{x}_1\boldsymbol{x}_n \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \boldsymbol{x}_n\boldsymbol{x}_1 & \ldots & \boldsymbol{x}_n\boldsymbol{x}_n \end{vmatrix} \tag{2.1–12}$$

is called *Gram's determinant* or the *gramian* of the system $\boldsymbol{x}_\kappa$. Now we shall prove a fundamental statement:

A necessary and sufficient condition for the system of vectors $\boldsymbol{x}_\kappa$, $\kappa = 1, \ldots, n$, to be linearly dependent is the vanishing of their gramian.

Suppose that the system is linearly dependent. Then we can find numbers a^κ, $\kappa = 1, \ldots, n$, not all zero, such that

$$\boldsymbol{x}_\mu a^\mu = \boldsymbol{o}. \tag{2.1–13}$$

On multiplying the expressions on the left by $\boldsymbol{x}_\lambda$, $\lambda = 1, \ldots, n$, successively we obtain a set of n homogeneous equations

$$(\boldsymbol{x}_\lambda \boldsymbol{x}_\mu)a^\mu = 0, \qquad \lambda = 1, \ldots, n. \tag{2.1–14}$$

Since this system possesses a non-trivial solution, its determinant, which is the gramian (2.1–12), must vanish.

Suppose now that the determinant (2.1–12) is zero. Then the system (2.1–14) possesses a non-trivial solution and on multiplying the first equation by a^1, the second by a^2, etc., we find, after adding the results obtained in this way,

$$(\boldsymbol{x}_\lambda \boldsymbol{x}_\mu)a^\lambda a^\mu = 0$$

or

$$(\boldsymbol{x}_\lambda a^\lambda)(\boldsymbol{x}_\mu a^\mu) = |\boldsymbol{x}_\lambda a^\lambda|^2 = 0.$$

This implies, however, $\boldsymbol{x}_\lambda a^\lambda = \boldsymbol{o}$, i.e., the system $\boldsymbol{x}_\kappa$ is linearly dependent.

2.2 – Reciprocal bases

2.2.1 – The orthogonal projection of a vector

It is not difficult to define the projection of a vector on a finite dimensional subspace $\mathfrak{R}_n$ of a metric vector space $\mathfrak{R}$. We shall prove the following statement:

Given a vector $\boldsymbol{x}$ there is in $\mathfrak{R}_n$ a uniquely determined vector $\mathsf{P}\boldsymbol{x}$ such that $\boldsymbol{x} - \mathsf{P}\boldsymbol{x}$ is orthogonal to all vectors of $\mathfrak{R}_n$.

More briefly stated: $\boldsymbol{x} - \mathsf{P}\boldsymbol{x}$ is orthogonal to $\mathfrak{R}_n$. First we wish to show that a vector of this kind is unique. Suppose we have two decompositions

$$x = y + (x-y) = y^* + (x-y^*)$$

such that y and y^* are in $\mathfrak{R}_n$ and $x-y$ and $x-y^*$ are orthogonal to $\mathfrak{R}_n$. It follows that

$$y-y^* = (x-y^*)-(x-y)$$

is orthogonal to all vectors of $\mathfrak{R}_n$. But $y-y^*$ is included in $\mathfrak{R}_n$. Hence $(y-y^*)(y-y^*) = 0$, that is $y-y^* = o$ or $y = y^*$.

Next we take in $\mathfrak{R}_n$ a basis x_κ, $\kappa = 1, \ldots, n$, and consider a vector of the type

$$x - x_\mu a^\mu. \tag{2.2–1}$$

The numbers a^μ, $\mu = 1, \ldots, n$, have to satisfy the condition that this vector is orthogonal to $\mathfrak{R}_n$, that is to say, to all vectors x_κ, $\kappa = 1, \ldots, n$. This condition gives rise to the system of inhomogeneous equations

$$(x_\lambda x_\mu) a^\mu = x_\lambda x, \qquad \lambda = 1, \ldots, n. \tag{2.2–2}$$

The determinant of the system is the gramian of the basis and according to the theorem of the previous section it is different from zero. Hence we can solve the system (2.2–2) and we obtain a vector

$$\mathsf{P}x = x_\mu a^\mu$$

such that $x - \mathsf{P}x$ is orthogonal to $\mathfrak{R}_n$. This proves the assertion.

The vector $\mathsf{P}x$ is called the *(orthogonal) projection* of x on $\mathfrak{R}_n$.

2.2.2 – Reciprocal bases

The following theorem establishes the fact that to a given basis x_κ in a finite dimensional space corresponds a second basis x^κ in a peculiar manner:

To a given basis

$$x_\kappa, \qquad \kappa = 1, \ldots, n, \tag{2.2–3}$$

of an n-dimensional metric space $\mathfrak{R}_n$ *corresponds a uniquely determined second basis*

$$x^\kappa, \qquad \kappa = 1, \ldots, n, \tag{2.2–4}$$

such that

$$\boxed{x_\lambda x^\mu = \delta^\mu_\lambda,} \qquad \lambda, \mu = 1, \ldots n. \tag{2.2–5}$$

Conversely, on account of (2.2–5), *the basis* (2.2–3) *is uniquely determined by* (2.2–4).

Omitting a vector x_λ from the basis (2.2–3) we obtain a set of $n-1$ vectors spanning an $(n-1)$-dimensional space. Let P_λ denote the operator performing the projection of a vector on this space. The vector

$$y = x_\lambda - \mathsf{P}_\lambda x_\lambda$$

is orthogonal to all vectors x_μ, $\mu \neq \lambda$, while $\mathsf{P}_\lambda x_\lambda$ is linearly dependent on these vectors. Hence $y \neq o$ and

$$0 < yy = yx_\lambda,$$

for $y(\mathsf{P}_\lambda x_\lambda) = 0$. On multiplying y by a suitable constant we find a vector y^* such that $y^* x_\lambda = 1$. Denoting this vector by x^λ we have obtained the desired result.

It is readily seen that the vectors (2.2–4) satisfying (2.2–5) are linearly independent. For, suppose that $x^\mu a_\mu = o$, where a_μ are numbers denoted by letters with subscripts in order to allow the application of the summation convention. But then $0 = x_\lambda(x^\mu a_\mu) = (x_\lambda x^\mu)a_\mu = \delta_\lambda^\mu a_\mu = a_\lambda$, $\lambda = 1, \ldots, n$. To conclude the proof we observe that we can proceed in the inverse direction starting with the basis x^κ and then we precisely obtain the original basis.

Either basis (2.2–3) or (2.2–4) is called the *reciprocal* of the other. The basis x_κ together with its reciprocal basis will be referred to as a *frame.* A frame will often be denoted by the symbol (κ) exhibiting the type of indices used. Thus another frame may be designated by (κ'), etc.

Finally we wish to refer a basis to its reciprocal. We may write

$$\boxed{x^\mu = x_\lambda g^{\lambda\mu},} \qquad (2.2\text{–}6)$$

where the $g^{\lambda\mu}$, $\lambda, \mu = 1, \ldots, n$ are the components of the vector x^μ with respect to the basis x_κ. Multiplying both members of (2.2–6) by x^κ, we find in view of (2.2–5):

$$x^\kappa x^\mu = x^\kappa x_\lambda g^{\lambda\mu} = \delta_\lambda^\kappa g^{\lambda\mu} = g^{\kappa\mu},$$

i.e.,

$$\boxed{g^{\lambda\mu} = x^\lambda x^\mu.} \qquad (2.2\text{–}7)$$

But we may also write

$$\boxed{x_\lambda = x^\mu g_{\lambda\mu}} \qquad (2.2\text{–}8)$$

and on multiplying by x_κ we isolate $g_{\lambda\kappa}$. Hence also

$$\boxed{g_{\lambda\mu} = x_\lambda x_\mu.} \qquad (2.2\text{–}9)$$

It is clear that $g_{\lambda\mu} = g_{\mu\lambda}$ and $g^{\lambda\mu} = g^{\mu\lambda}$.

It is also clear that the gramian of the basis $\boldsymbol{x}_\kappa$ may be written as

$$\det [g_{\lambda\mu}] \tag{2.2–10}$$

and that of the basis $\boldsymbol{x}^\kappa$ as

$$\det [g^{\lambda\mu}]. \tag{2.2–11}$$

A simple relation exists between (2.2–10) and (2.2–11). In fact, from (2.2–5) and (2.2–6) we deduce

$$\delta^\mu_\lambda = \boldsymbol{x}_\lambda \boldsymbol{x}^\mu = \boldsymbol{x}_\lambda \boldsymbol{x}_\kappa g^{\kappa\mu}$$

whence in view of (2.2–9)

$$\boxed{g_{\lambda\kappa} g^{\kappa\mu} = \delta^\mu_\lambda.} \tag{2.2–14}$$

On applying the multiplication rule for determinants we may infer that

$$\det [g_{\lambda\mu}] \det [g^{\lambda\mu}] = 1, \tag{2.2–15}$$

is the desired relation.

2.2.3 – Contravariant and covariant components of a vector

With reference to a frame, that is, a pair of reciprocal bases, a vector may be characterized by two kinds of components. That is, we may write

$$\boldsymbol{a} = \boldsymbol{x}_\kappa a^\kappa, \tag{2.2–16}$$

but also

$$\boldsymbol{a} = \boldsymbol{x}^\kappa a_\kappa. \tag{2.2–17}$$

The components in the equation (2.2–16) are called the *contravariant components* of $\boldsymbol{a}$ and the components occurring in (2.2–17) the *covariant components* with respect to the given frame. These names will be explained when we deal with the transformation of coordinates in the next section.

In view of (2.2–6) and (2.2–8) it is clear that

$$\boxed{a^\mu = a_\lambda g^{\lambda\mu}} \tag{2.2–18}$$

and

$$\boxed{a_\lambda = a^\mu g_{\lambda\mu}.} \tag{2.2–19}$$

The equations (2.2–18) and (2.2–19) describe a process which may be called the *raising* and the *lowering* of an index.

It is easy to find expressions for the inner product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in terms of their respective components. Let $\boldsymbol{b}$ denote the vector

$$\boldsymbol{b} = \boldsymbol{x}_\kappa b^\kappa = \boldsymbol{x}^\kappa b_\kappa. \tag{2.2–20}$$

Then, taking account of the rule (2.1–5)

$$\begin{aligned}\boldsymbol{ab} &= (\boldsymbol{x}_\lambda a^\lambda)(\boldsymbol{x}_\mu b^\mu) = (\boldsymbol{x}_\lambda \boldsymbol{x}_\mu) a^\lambda b^\mu \\ &= (\boldsymbol{x}_\lambda a^\lambda)(\boldsymbol{x}^\mu b_\mu) = (\boldsymbol{x}_\lambda \boldsymbol{x}^\mu) a^\lambda b_\mu \\ &= (\boldsymbol{x}^\lambda a_\lambda)(\boldsymbol{x}^\mu b_\mu) = (\boldsymbol{x}^\lambda \boldsymbol{x}^\mu) a_\lambda b_\mu.\end{aligned}$$

Hence, in view of (2.2–9), (2.2–5) and (2.2–7),

$$\boxed{\boldsymbol{ab} = g_{\lambda\mu} a^\lambda b^\mu = a^\kappa b_\kappa = g^{\lambda\mu} a_\lambda b_\mu.} \tag{2.2–21}$$

2.2.4 – Change of frame

Suppose we are given a second basis $\boldsymbol{x}_{\kappa'}$ related to the basis $\boldsymbol{x}_\kappa$ by means of the equations (1.1–11) or (1.1–12). Next we introduce the vectors

$$\boxed{\boldsymbol{x}^{\kappa'} = \boldsymbol{x}^\kappa q_\kappa^{\kappa'},} \qquad \kappa' = 1', \ldots, n', \tag{2.2–22}$$

where $\boldsymbol{x}^\kappa$, $\kappa = 1, \ldots, n$, are the elements of the basis reciprocal to the basis $\boldsymbol{x}_\kappa$. In view of (1.1–11) and (2.2–5) we have, taking account of (1.1–15) and (1.1–16):

$$\boldsymbol{x}_{\lambda'} \boldsymbol{x}^{\mu'} = \boldsymbol{x}_\lambda \boldsymbol{x}^\mu q_{\lambda'}^\lambda q_\mu^{\mu'} = \delta_\lambda^\mu q_{\lambda'}^\lambda q_\mu^{\mu'} = q_{\lambda'}^\kappa q_\kappa^{\mu'} = \delta_{\lambda'}^{\mu'}.$$

Hence the vectors $\boldsymbol{x}^{\kappa'}$ are the elements of a basis reciprocal to the basis $\boldsymbol{x}_\kappa$. The bases $\boldsymbol{x}_{\kappa'}$ and $\boldsymbol{x}^{\kappa'}$ constitute the frame (κ').

It is immediately seen that in addition to (1.1–18) we also have

$$\boxed{a_{\kappa'} = a_\kappa q_{\kappa'}^\kappa.} \tag{2.2–23}$$

Thus we see that the covariant components of a vector transform like the elements of a basis $\boldsymbol{x}_\kappa$, while the contravariant components transform like the elements of the reciprocal basis $\boldsymbol{x}^\kappa$.

2.3 – Orthonormal frames

2.3.1 – The Gram-Schmidt orthogonalization process

A set of vectors in a metric space $\mathfrak{R}$ is said to be *orthonormal* if for any pair of vectors $\boldsymbol{x}, \boldsymbol{y}$ included in the set

$$\boldsymbol{xy} = \begin{cases} 1, & \text{when } \boldsymbol{x} = \boldsymbol{y}, \\ 0, & \text{when } \boldsymbol{x} \neq \boldsymbol{y}. \end{cases} \tag{2.3–1}$$

An orthonormal system is always linearly independent.

In order to prove this statement we first observe that on account of (2.3–1) none of the vectors of the system is zero. Linear dependence means that there is a vector $\boldsymbol{x}$ which belongs to a space spanned by a finite number of vectors of the system distinct from $\boldsymbol{x}$. But since $\boldsymbol{x}$ is orthogonal to these vectors it is orthogonal to itself and is therefore the zero vector. This is in contradiction to the remark made above.

Next we consider a finite or infinite sequence

$$\boldsymbol{x}_\mu, \qquad \mu = 1, 2, \ldots \tag{2.3–2}$$

in a metric space $\mathfrak{R}$ and we suppose that this sequence is linearly independent. We assert:

There is an orthogonal sequence

$$\underset{h}{\boldsymbol{e}}, \qquad h = 1, 2, \ldots, \tag{2.3–3}$$

such that the space spanned by the vectors $\underset{h}{\boldsymbol{e}}$, $h = 1, \ldots, \mathrm{m}$, *coincides with the space spanned by* $\boldsymbol{x}_\mu$, $\mu = 1, \ldots, m$.

In order to emphasize the peculiar character of the orthonormal set we place the indices under the letters indicating the vectors.

Since $\boldsymbol{x}_1$ is linearly independent it is not the zero vector and we can normalize it by taking

$$\underset{1}{\boldsymbol{e}} = \boldsymbol{x}_1 \xi,$$

with $\xi = |\boldsymbol{x}_1|^{-1}$ (or $-|\boldsymbol{x}_1|^{-1}$, but this is of no importance). It is clear that $\boldsymbol{x}_1$ and $\underset{1}{\boldsymbol{e}}$ span the same space.

Suppose that we are already in possession of an orthonormal system

$$\underset{1}{\boldsymbol{e}}, \ldots, \underset{\mathrm{m}-1}{\boldsymbol{e}}$$

such that the vectors

$$\underset{h}{\boldsymbol{e}}, \qquad h = 1, \ldots, \mathrm{m}-1,$$

span the same space as $\boldsymbol{x}_\mu$, $\mu = 1, \ldots, m-1$. Next we introduce the vector

$$\boldsymbol{x} = \boldsymbol{x}_m - (\underset{1}{\boldsymbol{e}}\underset{1}{\xi} + \ldots + \underset{\mathrm{m}-1}{\boldsymbol{e}}\underset{\mathrm{m}-1}{\xi}),$$

more briefly written as

$$\boldsymbol{x} = \boldsymbol{x}_m - \sum_{h=1}^{\mathrm{m}-1} \underset{h}{\boldsymbol{e}}\underset{h}{\xi}.$$

When the numbers $\underset{h}{\xi}$ satisfy the conditions

$$\underset{h}{\boldsymbol{e}}\boldsymbol{x} = 0, \qquad h = 1, \ldots, \mathrm{m}-1,$$

then, evidently,

$$\underset{h}{\xi} = \underset{h}{e}\, x_m, \qquad h = 1, \ldots, m-1.$$

Since $\underset{1}{e}, \ldots, \underset{m-1}{e}$ span the same space as $x_1, \ldots, x_{m-1}$ it follows that x is a linear combination of $x_1, \ldots, x_m$. The vector x cannot be the zero vector, for x_m is not included in the space spanned by $\underset{1}{e}, \ldots, \underset{m-1}{e}$. Hence we can introduce the vector

$$\underset{m}{e} = x\xi$$

with $\xi = |x|^{-1}$ (or $\xi = -|x|^{-1}$, but that is irrelevant). Thus we extend the orthonormal set $\underset{1}{e}, \ldots, \underset{m-1}{e}$ to an orthonormal set $\underset{1}{e}, \ldots, \underset{m}{e}$.

Proceeding in this way we can make the orthonormal sequence as long as we desire, provided the original sequence is infinite. If, however, the original sequence is finite, the process — the *Gram-Schmidt orthogonalization process* — comes to an end after a finite number of steps.

2.3.2 – The orthonormal frame

A direct consequence of the result obtained in the previous section is: *A finite dimensional metric vector space possesses an orthonormal basis.*

The case $n = 0$ is of no interest. Under the assumption $n > 0$ we can find a basis consisting of linearly independent vectors and on applying the Gram-Schmidt process we can replace it by an orthonormal basis.

It is clear that the reciprocal of an orthonormal basis coincides with the original basis. As a consequence an orthonormal basis may also be referred to as an *orthonormal frame.*

For a subsequent application we need the following theorem:

The set of all vectors of an n-dimensional metric vector space $\mathfrak{R}_n$ orthogonal to a non-zero vector x of $\mathfrak{R}_n$ is an $(n-1)$-dimensional space $\mathfrak{R}_{n-1}$.

Let x_κ, $\kappa = 1, \ldots, n$, denote a basis of $\mathfrak{R}_n$ including the given vector x. By a suitable enumeration we may suppose that $x = x_1$. On applying the Gram-Schmidt process to the system x_κ we obtain a basis $\underset{h}{e}$, $h = 1, \ldots, n$, such that x_1 and $\underset{1}{e}$ span the same ray, while $\underset{2}{e}, \ldots, \underset{n}{e}$ span a space of dimension $n-1$ orthogonal to $\underset{1}{e}$, that is to x. It is easily checked that every vector orthogonal to x is included in this space $\mathfrak{R}_{n-1}$. This proves the assertion.

2.3.3 – Orthogonal components

The components of a vector with respect to an orthonormal frame are termed the *orthogonal* or *rectangular components* of the vector. We wish to list some simple expressions involving these components. The vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ may be represented as

$$\boldsymbol{a} = \sum_{h=1}^{n} \boldsymbol{e}_h a_h, \qquad \boldsymbol{b} = \sum_{h=1}^{n} \boldsymbol{e}_h b_h. \tag{2.3-4}$$

The inner product of these vectors is

$$\boldsymbol{ab} = \sum_{h=1}^{n} \sum_{k=1}^{n} (\boldsymbol{e}_h \boldsymbol{e}_k) a_h a_k,$$

But

$$\boldsymbol{e}_h \boldsymbol{e}_k = \begin{cases} 1, & \text{when } h = k, \\ 0, & \text{when } h \neq k. \end{cases} \tag{2.3-5}$$

Hence

$$\boldsymbol{ab} = \sum_{h=1}^{n} a_h b_h. \tag{2.3-6}$$

Next we consider a unit vector $\boldsymbol{e}$ and we denote by c_h, $h = 1, \ldots, n$, its rectangular components, that is,

$$\boldsymbol{e} = \sum_{h=1}^{n} \boldsymbol{e}_h c_h. \tag{2.3-7}$$

When θ_h denotes the angle between $\boldsymbol{e}$ and $\boldsymbol{e}_h$ we have

$$\cos \theta_h = \boldsymbol{e}\,\boldsymbol{e}_h, \qquad h = 1, \ldots, n.$$

On multiplying both members of (2.3–7) by $\boldsymbol{e}_k$ we find

$$c_k = \boldsymbol{e}\,\boldsymbol{e}_k \tag{2.3-9}$$

and thus it appears that

$$c_h = \cos \theta_h, \qquad h = 1, \cdots, n. \tag{2.3-10}$$

Consequently

$$\boxed{\boldsymbol{e} = \sum_{h=1}^{n} \boldsymbol{e}_h \cos \theta_h.} \tag{2.3-11}$$

By virtue of (2.2–6) and the fact that $\boldsymbol{ee} = 1$ we also have

$$\boxed{\sum_{h=1}^{n} \cos^2 \theta_h = 1.} \tag{2.3-12}$$

To conclude this section we wish to derive some useful expressions for the numbers $g_{\lambda\mu}$ and $g^{\lambda\mu}$ introduced in section 2.2.2.

In $\mathfrak{R}_n$ we consider a frame (κ) consisting of the vectors $\boldsymbol{x}_\kappa$ and $\boldsymbol{x}^\kappa$, $\kappa = 1, \ldots, n$. With reference to an orthonormal frame we may represent these vectors as

$$\boldsymbol{x}_\kappa = \sum_{h=1}^{n} \underset{h}{\boldsymbol{e}}\, \underset{h}{\xi_\kappa} \tag{2.3–13}$$

and

$$\boldsymbol{x}^\kappa = \sum_{h=1}^{n} \underset{h}{\boldsymbol{e}}\, \underset{h}{\xi^\kappa}, \qquad \kappa = 1, \cdots, n. \tag{2.3–14}$$

On the other hand we have

$$\underset{h}{\boldsymbol{e}} = \boldsymbol{x}_\kappa \underset{h}{e^\kappa} = \boldsymbol{x}^\kappa \underset{h}{e_\kappa}, \qquad h = 1, \ldots, \mathrm{n}, \tag{2.3–15}$$

where $\underset{h}{e^\kappa}$ and $\underset{h}{e_\kappa}$ are respectively the contravariant and covariant components of the $\underset{h}{\boldsymbol{e}}$ with respect to the frame (κ). On multiplying both members of (2.3–13) and (2.3–14) by $\underset{k}{\boldsymbol{e}}$ we get

$$\boldsymbol{x}_\kappa \underset{k}{\boldsymbol{e}} = \underset{k}{\xi_\kappa}, \qquad \boldsymbol{x}^\kappa \underset{k}{\boldsymbol{e}} = \underset{k}{\xi^\kappa}, \qquad k = 1, \ldots, \mathrm{n}, \quad \kappa = 1, \ldots, n.$$

From (2.3–15) we obtain by a similar process

$$\underset{h}{\boldsymbol{e}}\boldsymbol{x}^\lambda = \delta^\lambda_\kappa \underset{h}{e^\kappa} = \underset{h}{e^\lambda}, \qquad \underset{h}{\boldsymbol{e}}\boldsymbol{x}_\lambda = \delta^\kappa_\lambda \underset{h}{e_\kappa} = \underset{h}{e_\lambda}.$$

Hence

$$\underset{h}{\xi_\kappa} = \underset{h}{e_\kappa}, \qquad \underset{h}{\xi^\kappa} = \underset{h}{e^\kappa}$$

and thus it appears that

$$\boldsymbol{x}_\kappa = \sum_{h=1}^{n} \underset{h}{\boldsymbol{e}}\, \underset{h}{e_\kappa}, \qquad \boldsymbol{x}^\kappa = \sum_{h=1}^{n} \underset{h}{\boldsymbol{e}}\, \underset{h}{e^\kappa}. \tag{2.3–16}$$

Now it follows from (2.3–9) and (2.2–7) by means of the formula (2.3–6)

$$\boxed{g_{\lambda\mu} = \sum_{h=1}^{n} \underset{h}{e_\lambda}\, \underset{h}{e_\mu}} \tag{2.3–17}$$

and

$$\boxed{g^{\lambda\mu} = \sum_{h=1}^{n} \underset{h}{e^\lambda}\, \underset{h}{e^\mu}.} \tag{2.3–18}$$

As an application of (2.3–17) we shall prove:

The gramian of a linearly independent set of vectors is always positive.

Let $\boldsymbol{x}_\kappa$, $\kappa = 1, \ldots, n$, be this set. In the space spanned by it we introduce an orthonormal frame. Then

$$\det\,[g_{\lambda\mu}] = \det\,[\sum_{h=1}^{n} \underset{h}{e_\lambda}\, \underset{h}{e_\mu}] = \det{}^2\,[\underset{h}{e_\lambda}] > 0.$$

It should be noticed that this theorem, combined with the theorem of section 2.1.3, reduces to Schwarz's inequality when $n = 2$.

CHAPTER 3

BILINEAR AND QUADRATIC FORMS

With an operator in a metric vector space is associated a real-valued function of a pair of vector variables which possesses many interesting properties. These properties also find a wide field of applications beyond the frontiers of differential geometry.

3.1 – Bilinear forms

3.1.1 – THE BILINEAR FORM ASSOCIATED WITH A LINEAR OPERATOR

Let A denote a linear operator in a metric vector space $\mathfrak{R}$. If $\boldsymbol{x}$ and $\boldsymbol{y}$ are two vectors we may consider the inner product of the vectors $\boldsymbol{x}$ and $\mathsf{A}\boldsymbol{y}$

$$\boldsymbol{x}\mathsf{A}\boldsymbol{y} \tag{3.1–1}$$

This expression is evidently a function of two vector variables $\boldsymbol{x}$ and $\boldsymbol{y}$. It is named a *bilinear form* on account of the fact that it behaves linearly with respect to each variable separately. That is to say

$$\boldsymbol{x}\mathsf{A}(\boldsymbol{u}a+\boldsymbol{v}b) = \boldsymbol{x}(\mathsf{A}\boldsymbol{u}a+\mathsf{A}\boldsymbol{v}b) = \boldsymbol{x}\mathsf{A}\boldsymbol{u}a+\boldsymbol{x}\mathsf{A}\boldsymbol{v}b$$

and

$$(\boldsymbol{u}a+\boldsymbol{v}b)\mathsf{A}\boldsymbol{y} = \boldsymbol{u}\mathsf{A}\boldsymbol{y}a+\boldsymbol{v}\mathsf{A}\boldsymbol{y}b,$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are arbitrary vectors and a and b arbitrary numbers.

In general the forms $\boldsymbol{x}\mathsf{A}\boldsymbol{y}$ and $\boldsymbol{y}\mathsf{A}\boldsymbol{x}$ are not the same, for in the definition of (3.1–1) the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ do not play the same part. An operator A, however, satisfying the condition

$$\boldsymbol{x}\mathsf{A}\boldsymbol{y} = \boldsymbol{y}\mathsf{A}\boldsymbol{x} \tag{3.1–2}$$

for every pair of vectors $\boldsymbol{x}, \boldsymbol{y}$ is called *symmetric*. This name is also adopted for the associated bilinear form.

In the particular case that A is the identity the form (3.1–1) coincides with the inner product $\boldsymbol{x}\boldsymbol{y}$ which is evidently symmetric.

The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be *conjugate* with respect to the symmetric operator A when

$$\boldsymbol{x}\mathsf{A}\boldsymbol{y} = \boldsymbol{y}\mathsf{A}\boldsymbol{x} = 0. \tag{3.1–3}$$

This property is a generalization of the concept of orthogonality, for vectors conjugate with respect to the identity are orthogonal.

A vector y is conjugate to all vectors of the space if and only if the vector is singular for the operator.

For, if y is singular then $\mathsf{A}y = o$ and thus $x\mathsf{A}y = 0$ for all vectors of the space. If, conversely, $x\mathsf{A}y = 0$ for every vector x then when $x = \mathsf{A}y$, $(\mathsf{A}y)\mathsf{A}y = 0$ or $\mathsf{A}y = o$. Hence y is singular.

3.1.2 – Bilinear forms in a finite dimensional metric space

Suppose that $\mathfrak{R}_n$ is an n-dimensional metric space referred to a frame (κ). In section 1.2.5 we introduced the components of the operator A by means of the equations (1.2–18). It is readily verified that these components may be written as

$$a^\lambda{}_\mu = x^\lambda \mathsf{A} x_\mu. \tag{3.1–4}$$

This suggests to us the idea of defining other components in the following way:

$$a_\lambda{}^\mu = x_\lambda \mathsf{A} x^\mu, \tag{3.1–5}$$

$$a_{\lambda\mu} = x_\lambda \mathsf{A} x_\mu, \tag{3.1–6}$$

and

$$a^{\lambda\mu} = x^\lambda \mathsf{A} x^\mu. \tag{3.1–7}$$

In terms of the various types of components of x, y with respect to (κ) we evidently have

$$x\mathsf{A}y = a^\lambda{}_\mu x_\lambda y^\mu = a_\lambda{}^\mu x^\lambda y_\mu = a_{\lambda\mu} x^\lambda y^\mu = a^{\lambda\mu} x_\lambda y_\mu. \tag{3.1–8}$$

The symmetry of A is expressed by the equations

$$a_{\lambda\mu} = a_{\mu\lambda}, \qquad a^{\lambda\mu} = a^{\mu\lambda}, \quad a_\lambda{}^\mu = a^\mu{}_\lambda \tag{3.1–9}$$

Either of these equations implies the remaining two.

In the case of a symmetric operator we shall write a^λ_μ instead of $a^\lambda{}_\mu$, in accordance with the last equation (3.1–9). This has already been anticipated by the notation δ^λ_μ for the Kronecker symbol representing the identity.

The components $a_{\lambda\mu}$ are called *the purely covariant components* of the operator A (or of the associated bilinear form). The components $a^{\lambda\mu}$ are called the *purely contravariant components*, while $a^\lambda{}_\mu$ and $a_\lambda{}^\mu$ are referred to as the *mixed* components.

In particular $g_{\lambda\mu}$ and $g^{\lambda\mu}$ are respectively the purely covariant and purely contravariant components of the identity.

3.1.3 – The equations of transformation

The components of an operator A with respect to a second frame (κ') are related to the components with respect to a frame (κ) in a simple manner.

We already encountered a relation of this kind in section 1.2.6. Similar equations may be derived at once from the expressions (3.1–5), (3.1–6) and (3.1–7), taking the equations (1.1–11) and (2.2–22) into account. Because of the linearity we have

$$a_{\lambda'\mu'} = a_{\lambda\mu} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} \tag{3.1–10}$$

and

$$a^{\lambda'\mu'} = a^{\lambda\mu} q^{\lambda'}_{\lambda} q^{\mu'}_{\mu}, \tag{3.1–11}$$

while $a_{\lambda}{}^{\mu}$ transforms like $a^{\mu}{}_{\lambda}$.

If, conversely, we let correspond to every frame (κ) *a system of numbers* $a_{\lambda\mu}$ *in such a way that the numbers corresponding to different frames are related as in the equations* (3.1–10), *then they uniquely define an operator* A *having these numbers as components with respect to the given frame.*

We observe that the numbers

$$u_{\lambda} = a_{\lambda\mu} y^{\mu} \tag{3.1–12}$$

may be considered as the components of a vector $\boldsymbol{u}$. In fact, with reference to a system (κ') we have

$$u_{\lambda'} = a_{\lambda'\mu'} y^{\mu'} = a_{\lambda\mu} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} y^{\mu'} = a_{\lambda\mu} q^{\lambda}_{\lambda'} y^{\mu} = u_{\lambda} q^{\lambda}_{\lambda'}$$

and it is perfectly clear that the correspondence between $\boldsymbol{y}$ and $\boldsymbol{u}$ is that of a linear mapping.

Any statements about the components are significant only when they remain valid on replacing the components by those corresponding to another frame. Thus, for instance, a statement about symmetry need only be verified in one frame. For, when $a_{\lambda\mu} = a_{\mu\lambda}$ in a frame (κ) then

$$a_{\lambda'\mu'} = a_{\lambda\mu} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} = a_{\mu\lambda} q^{\mu}_{\mu'} q^{\lambda}_{\lambda'} = a_{\mu'\lambda'}$$

in the frame (κ').

It is sufficient to give the components of a certain type. In fact, from (2.2–6) and (2.2–8) we deduce with reference to (3.1–4)—(3.1–7):

$$a_{\lambda\mu} = g_{\lambda\kappa} a^{\kappa}{}_{\mu} = g_{\kappa\mu} a_{\lambda}{}^{\kappa}, \tag{3.1–13}$$

$$a^{\lambda\mu} = g^{\lambda\kappa} a_{\kappa}{}^{\mu} = g^{\kappa\mu} a^{\lambda}{}_{\kappa}, \tag{3.1–14}$$

$$a^{\lambda}{}_{\mu} = g^{\lambda\kappa} a_{\kappa\mu} = g_{\kappa\mu} a^{\lambda\kappa}, \tag{3.1–15}$$

$$a_{\lambda}{}^{\mu} = g_{\lambda\kappa} a^{\kappa\mu} = g^{\kappa\mu} a_{\lambda\kappa}. \tag{3.1–16}$$

These relations are actually invariant with respect to a change of frame. Thus, for instance,

$$\begin{aligned} a_{\lambda'\mu'} &= g_{\lambda'\kappa'} a^{\kappa'}{}_{\mu'} = g_{\lambda\kappa} a^{\rho}{}_{\mu} q^{\lambda}_{\lambda'} q^{\kappa}_{\kappa'} q^{\kappa'}_{\rho} q^{\mu'}_{\mu} \\ &= g_{\lambda\kappa} a^{\rho}{}_{\mu} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} \delta^{\kappa}_{\rho} = g_{\lambda\kappa} a^{\kappa}{}_{\mu} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} = a_{\lambda\mu} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'}, \end{aligned}$$

that is to say, the formation of $a_{\lambda\mu}$ from $a^{\lambda}{}_{\mu}$ is accomplished in different frames in exactly the same way.

3.2 – Eigenvalues and eigenvectors of a symmetric linear operator

3.2.1 – Symmetric operators in metric space

In section 1.2.7 we introduced the concept of eigenvalue and eigenvector of a linear operator. We now wish to study the properties of these concepts more closely in the case where the operator is defined throughout a metric space and is symmetric. First we shall establish the theorem:

Two eigenvectors of a symmetric linear operator corresponding to different eigenvalues are orthogonal and at the same time conjugate with respect to the operator.

Let

$$\mathsf{A}\boldsymbol{x} = \sigma\boldsymbol{x}, \qquad \mathsf{A}\boldsymbol{y} = \tau\boldsymbol{y}, \qquad \sigma \neq \tau.$$

Then

$$\boldsymbol{x}\mathsf{A}\boldsymbol{y} = \boldsymbol{x}\tau\boldsymbol{y} = \tau\boldsymbol{x}\boldsymbol{y}, \qquad \boldsymbol{y}\mathsf{A}\boldsymbol{x} = \boldsymbol{y}\sigma\boldsymbol{x} = \sigma\boldsymbol{x}\boldsymbol{y}.$$

Since A is symmetric we infer that $\sigma\boldsymbol{x}\boldsymbol{y} = \tau\boldsymbol{x}\boldsymbol{y}$ or $(\sigma-\tau)\boldsymbol{x}\boldsymbol{y} = 0$. Hence $\boldsymbol{x}\boldsymbol{y} = 0$ and $\boldsymbol{x}\mathsf{A}\boldsymbol{y} = \tau\boldsymbol{x}\boldsymbol{y} = 0$.

Another useful theorem states:

A symmetric linear operator sends every vector orthogonal to an eigenvector into a vector which is again orthogonal to this eigenvector.

Let $\boldsymbol{x}$ denote an eigenvector belonging to the eigenvalue σ. If $\boldsymbol{y}$ is orthogonal to $\boldsymbol{x}$ then

$$\boldsymbol{x}\mathsf{A}\boldsymbol{y} = \boldsymbol{y}\mathsf{A}\boldsymbol{x} = \boldsymbol{x}\sigma\boldsymbol{y} = \sigma\boldsymbol{y}\boldsymbol{x} = 0.$$

Hence $\mathsf{A}\boldsymbol{y}$ is orthogonal to $\boldsymbol{x}$.

The converse statement is also true:

If $\boldsymbol{x} \neq \boldsymbol{o}$ is a vector such that $\boldsymbol{y}\boldsymbol{x} = 0$ always implies $\boldsymbol{y}\mathsf{A}\boldsymbol{x} = 0$, where A is a symmetric operator, then $\boldsymbol{x}$ is an eigenvector of A.

If $\boldsymbol{x} \neq \boldsymbol{o}$ we can find a number σ such that

$$\boldsymbol{x}\mathsf{A}\boldsymbol{x} = \sigma\boldsymbol{x}\boldsymbol{x}.$$

This equation expresses the fact that the vectors $\boldsymbol{x}$ and $\boldsymbol{y} = \mathsf{A}\boldsymbol{x}-\sigma\boldsymbol{x}$ are orthogonal. But by hypothesis $\boldsymbol{y}\mathsf{A}\boldsymbol{x} = 0$ and therefore $\boldsymbol{y}\boldsymbol{y} = \boldsymbol{y}(\mathsf{A}\boldsymbol{x}-\sigma\boldsymbol{x}) = \boldsymbol{y}\mathsf{A}\boldsymbol{x}-\sigma\boldsymbol{y}\boldsymbol{x} = 0$. It follows that $\boldsymbol{y} = \boldsymbol{o}$, that is to say $\mathsf{A}\boldsymbol{x} = \sigma\boldsymbol{x}$. This proves the assertion.

3.2.2 – The roots of the characteristic equation

We pointed out in section 1.2.7 that the eigenvalues of an operator are among the roots of an algebraic equation of the type (1.2–38). Since the purely covariant components of the operator $\mathsf{A}-\sigma$ are $a_{\lambda\mu}-\sigma g_{\lambda\mu}$, the eigenvalues also satisfy

$$\det\,[a_{\lambda\mu}-\sigma g_{\lambda\mu}] = 0. \tag{3.2–1}$$

It is easy to verify that (3.2–1) arises from (1.2–38) by multiplication by $\det\,[g_{\lambda\mu}]$. Hence these equations yield the same roots.

Now we wish to prove the fundamental theorem:

A characteristic equation associated with a symmetric linear operator in a metric vector space has only real roots.

We know that in the case $n \geqq 1$ — and this is the only interesting case — the equation (3.2–1) has at least one root which may, however, be a non-real number. The corresponding system of equations

$$(a_{\lambda\mu}-\sigma g_{\lambda\mu})x^{\mu} = 0, \qquad \lambda = 1, \ldots, n, \tag{3.2–2}$$

then has a non-trivial solution x^{μ}, $\mu = 1, \ldots, n$, but these numbers need not be real, either. Since the numbers $a_{\lambda\mu}$ and $g_{\lambda\mu}$ are real, the conjugate complex numbers $\bar{x}^{\mu}$ satisfy the system

$$(a_{\lambda\mu}-\bar{\sigma} g_{\lambda\mu})\bar{x}^{\mu} = 0, \tag{3.2–3}$$

where $\bar{\sigma}$ is the conjugate complex of σ. From (3.2–2) we deduce

$$(a_{\lambda\mu}-\sigma g_{\lambda\mu})\bar{x}^{\lambda}x^{\mu} = 0 \tag{3.2–4}$$

and from (3.2–3)

$$(a_{\lambda\mu}-\bar{\sigma} g_{\lambda\mu})x^{\lambda}\bar{x}^{\mu} = 0. \tag{3.2–5}$$

The symmetry of A implies

$$a_{\lambda\mu}\bar{x}^{\lambda}x^{\mu} = a_{\mu\lambda}\bar{x}^{\lambda}x^{\mu} = a_{\lambda\mu}\bar{x}^{\mu}x^{\lambda} = a_{\lambda\mu}x^{\lambda}\bar{x}^{\mu}$$

and for a similar reason we also have

$$g_{\lambda\mu}\bar{x}^{\lambda}x^{\mu} = g_{\lambda\mu}x^{\lambda}\bar{x}^{\mu}.$$

On subtracting corresponding members of (3.2–4) and (3.2–5) we get

$$(\sigma-\bar{\sigma})g_{\lambda\mu}\bar{x}^{\lambda}x^{\mu} = 0. \tag{3.2–6}$$

Next we wish to show that $g_{\lambda\mu}\bar{x}^{\lambda}x^{\mu} \neq 0$. Let us put

$$x^{\lambda} = u^{\lambda}+\mathrm{i}v^{\lambda}$$

where u^{λ}, v^{λ} are real, $\lambda = 1, \ldots, n$. Then

$$\begin{aligned} g_{\lambda\mu}\bar{x}^{\lambda}x^{\mu} &= g_{\lambda\mu}(u^{\lambda}-\mathrm{i}v^{\lambda})(u^{\mu}+\mathrm{i}v^{\mu}) \\ &= g_{\lambda\mu}u^{\lambda}u^{\mu}-\mathrm{i}g_{\lambda\mu}v^{\lambda}u^{\mu}+\mathrm{i}g_{\lambda\mu}u^{\lambda}v^{\mu}+g_{\lambda\mu}v^{\lambda}v^{\mu} \\ &= g_{\lambda\mu}u^{\lambda}u^{\mu}+g_{\lambda\mu}v^{\lambda}v^{\mu} = \mathbf{uu}+\mathbf{vv}, \end{aligned}$$

where $u = x_\kappa u^\kappa$, $v = x_\kappa v^\kappa$. Because of the fact that not all numbers x^κ, $\kappa = 1, \ldots, n$, are equal to zero, at least one of the vectors u, v is not the zero vector. Hence $g_{\lambda\mu}\bar{x}^\lambda x^\mu > 0$ and it follows from (3.2–6) that $\sigma = \bar{\sigma}$, that is to say, σ is a real number.

3.2.3 – The principal directions

A vector of norm unity is often referred to as a *direction*. By a *principal direction* of an operator A in a metric vector space is understood a unit eigenvector of the operator. The main theorem of this section states:

A symmetric linear operator A *in an n-dimensional metric space* $\mathfrak{R}_n$ *possesses a system of n mutually orthogonal principal directions.*

The case $n = 0$ is of no interest. We assume, therefore, $n > 0$. There exists an eigenvalue $\underset{1}{\sigma}$, being a real root of the characteristic equation associated with the operator. Let $\underset{1}{p}$ be an eigenvector of norm unity beloging to $\underset{1}{\sigma}$. In section 2.3.2 we proved that all vectors of $\mathfrak{R}_n$ orthogonal to $\underset{1}{p}$ constitute a linear space $\mathfrak{R}_{n-1}$ of dimension $n-1$. By virtue of a theorem of section 3.2.1 the operator transforms $\mathfrak{R}_{n-1}$ into itself. Considered as an operator in $\mathfrak{R}_{n-1}$ it is again symmetric. If $n > 1$ we can repeat the previous reasoning, that is to say, we can find an eigenvalue $\underset{2}{\sigma}$ (not necessarily different from $\underset{1}{\sigma}$) and a unit eigenvector $\underset{2}{p}$ belonging to it and included in $\mathfrak{R}_{n-1}$. Proceeding in this way we obtain an orthogonal frame of vectors $\underset{h}{p}$, such that

$$\mathsf{A}\underset{h}{p} = \underset{h}{\sigma}\,\underset{h}{p}, \qquad h = 1, \ldots, \mathrm{n}. \tag{3.2–7}$$

When two eigenvectors belong to the same eigenvalue then any linear combination of these eigenvectors is either zero or is again an eigenvector belonging to that eigenvalue. Thus we arrive at the result that to an eigenvalue of multiplicity k — that is a root of multiplicity k of the characteristic equation — corresponds a k-dimensional eigenspace $\mathfrak{R}_k$. Any orthonormal frame of $\mathfrak{R}_k$ can be a part of the desired frame in $\mathfrak{R}_n$. Only in the case that all eigenvalues are different is the orthonormal system of principal directions uniquely determined.

With reference to a frame of principal directions the operator A is represented by the diagonal matrix

$$\begin{bmatrix} \underset{1}{\sigma} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \underset{\mathrm{n}}{\sigma} \end{bmatrix}, \tag{3.2–8}$$

all elements outside the diagonal being zero.

A particular case presents itself when all eigenvalues are equal to the same number σ. Then every vector $\boldsymbol{x} \neq \boldsymbol{o}$ is an eigenvector belonging to σ, since the whole space $\mathfrak{R}_n$ is an eigenspace. It is clear that

$$\mathsf{A} = \sigma 1, \tag{3.2–9}$$

where 1 denotes the identity. As a consequence the covariant components of A are

$$a_{\lambda\mu} = \sigma g_{\lambda\mu}. \tag{3.2–10}$$

A corollary of the main theorem states:

The rank r of a symmetric operator A *in an n-dimensional metric space is equal to the maximal number of linearly independent eigenvectors belonging to the non-vanishing eigenvalues.*

Let $\underset{h}{\boldsymbol{p}}$, $h = 1, \ldots, n$, denote the vectors of an orthonormal system of principal directions belonging to the eigenvalues $\underset{h}{\sigma}$ respectively. We arrange these eigenvectors in order such that $\underset{1}{\sigma}, \ldots, \underset{r}{\sigma}$ are all different from zero, whereas the remaining values vanish. An arbitrary vector $\boldsymbol{x}$ can be written as

$$\boldsymbol{x} = \underset{1}{\boldsymbol{p}}\underset{1}{x} + \ldots + \underset{r}{\boldsymbol{p}}\underset{r}{x} + \ldots + \underset{n}{\boldsymbol{p}}\underset{n}{x}$$

and it follows that

$$\mathsf{A}\boldsymbol{x} = \mathsf{A}\underset{1}{\boldsymbol{p}}\underset{1}{x} + \ldots + \mathsf{A}\underset{r}{\boldsymbol{p}}\underset{r}{x} + \ldots + \mathsf{A}\underset{n}{\boldsymbol{p}}\underset{n}{x} = \underset{1}{\boldsymbol{p}}\underset{1}{\sigma}\underset{1}{x} + \ldots + \underset{r}{\boldsymbol{p}}\underset{r}{\sigma}\underset{r}{x}.$$

Hence $\mathsf{A}\boldsymbol{x}$ is included in the space spanned by the vectors $\underset{1}{\boldsymbol{p}}, \ldots, \underset{r}{\boldsymbol{p}}$. On the other hand $\mathsf{A}\underset{1}{\boldsymbol{p}}, \ldots, \mathsf{A}\underset{r}{\boldsymbol{p}}$ are linearly independent. It follows that the space of all vectors $\mathsf{A}\boldsymbol{x}$ is exactly the space spanned by $\underset{1}{\boldsymbol{p}}, \ldots, \underset{r}{\boldsymbol{p}}$.

3.3 – Quadratic forms

3.3.1 – Euler's theorem

A *quadratic form*

$$\boldsymbol{x}\mathsf{A}\boldsymbol{x} \tag{3.3–1}$$

arises from (3.1–1) when we let $\boldsymbol{x}$ and $\boldsymbol{y}$ coincide. It is our aim to study the range of values of this form when $\boldsymbol{x}$ varies throughout the set of all unit vectors of $\mathfrak{R}_n$. It is clear that the value of (3.3–1) can be obtained from the value of this expression for a unit vector by multiplication by the norm of $\boldsymbol{x}$. Hence it is sufficient to restrict ourselves to the case that the norm of $\boldsymbol{x}$ is unity.

The following theorem, *Euler's theorem*, states that everything is known about the values of a quadratic form (3.3–1) when we know the eigenvalues of the operator A.

If the unit vector $\boldsymbol{e}$ *makes the angles* $\underset{h}{\theta}$ *with the vectors* $\underset{h}{\boldsymbol{p}}$, $h = 1, \ldots, \mathrm{n}$, *of an orthonormal frame of principal directions, then the value* σ *of the quadratic form* (3.3–1) *for* $\boldsymbol{x} = \boldsymbol{e}$ *is equal to*

$$\boxed{\sigma = \sum_{h=1}^{\mathrm{n}} \underset{h}{\sigma} \cos^2 \underset{h}{\theta}.} \tag{3.3–2}$$

As usual, $\underset{h}{\boldsymbol{p}}$ denotes an eigenvector of norm unity belonging to the eigenvalue $\underset{h}{\sigma}$. First we observe that $\underset{h}{\sigma}$ is exactly the value of (3.3–1) for $\boldsymbol{x} = \underset{h}{\boldsymbol{p}}$ For, from

$$\mathsf{A}\underset{h}{\boldsymbol{p}} = \underset{h}{\sigma}\underset{h}{\boldsymbol{p}}$$

it follows that

$$\underset{h}{\boldsymbol{p}}\mathsf{A}\underset{h}{\boldsymbol{p}} = \underset{h}{\sigma}\underset{h}{\boldsymbol{p}}\underset{h}{\boldsymbol{p}} = \underset{h}{\sigma}.$$

Taking account of (2.3–11) we find in view of (2.3–9) and (2.3–10)

$$\sigma = \boldsymbol{e}\mathsf{A}\boldsymbol{e} = \boldsymbol{e}\mathsf{A}\sum_{h=1}^{\mathrm{n}} \underset{h}{\boldsymbol{p}} \cos \underset{h}{\theta} = \sum_{h=1}^{\mathrm{n}} \underset{h}{\sigma}\boldsymbol{e}\underset{h}{\boldsymbol{p}} \cos \underset{h}{\theta} = \sum_{h=1}^{\mathrm{n}} \underset{h}{\sigma} \cos^2 \underset{h}{\theta}.$$

A consequence of Euler's theorem is:

The sum of the values of the quadratic form (3.3–1) *for n mutually orthogonal unit vectors is equal to the sum of the eigenvalues of the operator* A.

Let $\underset{h}{\boldsymbol{e}}$ denote any orthonormal frame and $\underset{h}{\boldsymbol{p}}$, $h = 1, \ldots, \mathrm{n}$, an orthonormal frame of principal directions. We put

$$\cos \underset{hk}{\theta} = \underset{h}{\boldsymbol{e}}\underset{k}{\boldsymbol{p}}, \qquad h, k = 1, \ldots, \mathrm{n}.$$

Referring to (2.3–12) we infer that

$$\sum_{h=1}^{\mathrm{n}} \cos^2 \underset{hk}{\theta} = 1 \qquad k = 1, \ldots, \mathrm{n},$$

and by virtue of (3.3–2) we have

$$\sum_{h=1}^{\mathrm{n}} \underset{h}{\boldsymbol{e}}\mathsf{A}\underset{h}{\boldsymbol{e}} = \sum_{h=1}^{\mathrm{n}}\sum_{k=1}^{\mathrm{n}} \underset{k}{\sigma} \cos^2 \underset{hk}{\theta} = \sum_{k=1}^{\mathrm{n}} \underset{k}{\sigma} \sum_{h=1}^{\mathrm{n}} \cos^2 \underset{hk}{\theta} = \sum_{k=1}^{\mathrm{n}} \underset{k}{\sigma},$$

the desired result.

3.3.2 – Some fundamental invariants

It is clear that any expression in terms of the eigenvalues gives rise to invariant expressions in terms of the components of the operator. We wish to derive two important formulas which are of fundamental importance for geometric applications.

First we consider the sum of the eigenvalues. Let $\underset{h}{\boldsymbol{p}}$ denote the vectors of an orthonormal frame of principal directions. The contravariant components of $\underset{h}{\boldsymbol{p}}$ with respect to an arbitrary frame (κ) are denoted as $\underset{h}{p}^{\kappa}$. Then

$$\sum_{h=1}^{n} \underset{h}{\sigma} = \sum_{h=1}^{n} \underset{h}{\boldsymbol{p}} \mathsf{A} \underset{h}{\boldsymbol{p}} = \sum_{h=1}^{n} a_{\lambda\mu} \underset{h}{p}^{\lambda} \underset{h}{p}^{\mu} = a_{\lambda\mu} \sum_{h=1}^{n} \underset{h}{p}^{\lambda} \underset{h}{p}^{\mu}.$$

Combining this result with that of (2.3–18) we get

$$\boxed{\sum_{h=1}^{n} \underset{h}{\sigma} = a_{\lambda\mu} g^{\lambda\mu} = a_{\kappa}^{\kappa}.} \tag{3.3–4}$$

A particular case is the formula

$$n = \delta_{\kappa}^{\kappa}, \tag{3.2–5}$$

for the eigenvalues of the identity are all unity. The formula (3.2–5) is, however, a direct consequence of the definition of the Kronecker symbol.

The number a_{κ}^{κ} is called the *trace* of the operator A. As we pointed out above it is an invariant with respect to change of frame. We may also verify this invariance by direct computation:

$$a_{\kappa'}^{\kappa'} = a_{\lambda}^{\kappa} q_{\kappa'}^{\lambda} q_{\kappa}^{\kappa'} = a_{\lambda}^{\kappa} \delta_{\kappa}^{\lambda} = a_{\kappa}^{\kappa}.$$

Another interesting invariant is found when we consider the iteration $\mathsf{A}^2 = \mathsf{AA}$ of the operator A. This operator is again symmetric as follows from

$$\boldsymbol{x}\mathsf{A}^2\boldsymbol{y} = \boldsymbol{x}\mathsf{A}(\mathsf{A}\boldsymbol{y}) = \mathsf{A}\boldsymbol{y}\mathsf{A}\boldsymbol{x} = \mathsf{A}\boldsymbol{x}\mathsf{A}\boldsymbol{y} = \boldsymbol{y}\mathsf{A}^2\boldsymbol{x}.$$

Now let $\boldsymbol{x}$ denote an eigenvector of the operator A. Then

$$\mathsf{A}\boldsymbol{x} = \sigma\boldsymbol{x}$$

and

$$\mathsf{A}^2\boldsymbol{x} = \mathsf{A}\sigma\boldsymbol{x} = \sigma\mathsf{A}\boldsymbol{x} = \sigma^2\boldsymbol{x},$$

and thus it appears that $\boldsymbol{x}$ is also an eigenvector of A^2, while the square of the corresponding eigenvalue of A is an eigenvalue of A^2. The components of A^2 with respect to a frame (κ) are

$$b_{\mu}^{\lambda} = a_{\rho}^{\lambda} a_{\mu}^{\rho}$$

and it follows from (3.3–4) that

$$\sum_{h=1}^{n} \underset{h}{\sigma}^2 = b_{\kappa}^{\kappa} = a_{\kappa}^{\lambda} a_{\lambda}^{\kappa}. \tag{3.3–5}$$

Without further explanation it is clear how we may find a sum of higher powers of the eigenvalues.

The results (3.3–5) and (3.3–4) enable us to obtain another important formula, observing that

$$2 \sum_{h<k} \underset{h}{\sigma} \underset{k}{\sigma} = \left(\sum_{h=1}^{n} \underset{h}{\sigma}\right)^2 - \sum_{h=1}^{n} \underset{h}{\sigma}^2.$$

Hence

$$\boxed{2 \sum_{h<k} \underset{h}{\sigma} \underset{k}{\sigma} = a_\lambda^\lambda a_\mu^\mu - a_\mu^\lambda a_\lambda^\mu.} \tag{3.3–6}$$

It is easily verified that this equation may be written in the form

$$2 \sum_{h<k} \underset{h}{\sigma} \underset{k}{\sigma} = -g^{\alpha\mu} g^{\beta\lambda} (a_{\alpha\lambda} a_{\beta\mu} - a_{\beta\lambda} a_{\alpha\mu}) \tag{3.3–7}$$

or

$$2 \sum_{h<k} \underset{h}{\sigma} \underset{h}{\sigma} = -g^{\alpha\mu} g^{\beta\lambda} \begin{vmatrix} a_{\alpha\lambda} & a_{\alpha\mu} \\ a_{\beta\lambda} & a_{\beta\mu} \end{vmatrix}. \tag{3.3–8}$$

3.3.3 – Extremal properties of the eigenvalues

Again let $\underset{h}{\boldsymbol{p}}$, $h = 1, \ldots, n$, denote a system of principal directions of the symmetric operator **A** in an n-dimensional metric space. From now on we shall assume that the corresponding eigenvalues are arranged in order such that

$$\underset{1}{\sigma} \geqq \underset{2}{\sigma} \geqq \ldots \geqq \underset{n}{\sigma} \tag{3.3–9}$$

First we state:

The eigenvalue $\underset{m}{\sigma}$, *is the maximal value of the quadratic form* (3.3–1) *for a unit vector* $\boldsymbol{e}$ *which satisfies the conditions*

$$\boldsymbol{e}\underset{1}{\boldsymbol{p}} = \ldots = \boldsymbol{e}\underset{m-1}{\boldsymbol{p}} = 0. \tag{3.3–10}$$

This means that in the case m = 1 that there are no conditions.

The proof is easy when we apply Euler's theorem. The angles $\underset{1}{\theta}, \ldots, \underset{m-1}{\theta}$ are equal to $\frac{1}{2}\pi$ and in view of (3.3–2) we have

$$\begin{aligned} \sigma = \boldsymbol{e}\mathsf{A}\boldsymbol{e} &= \underset{m}{\sigma} \cos^2 \underset{m}{\theta} + \ldots + \underset{n}{\sigma} \cos^2 \underset{n}{\theta} \leqq \underset{m}{\sigma}(\cos^2 \underset{m}{\theta} + \ldots + \cos^2 \underset{n}{\theta}) \\ &= \underset{m}{\sigma} \sum_{h=1}^{n} \cos^2 \underset{h}{\theta} = \underset{m}{\sigma}. \end{aligned}$$

Hence σ does not exceed $\underset{m}{\sigma}$. On the other hand $\underset{m}{\boldsymbol{p}}$ satisfies the conditions (3.3–10) and $\underset{m}{\sigma}$ is precisely the value of (3.3–1) for $\boldsymbol{x} = \underset{m}{\boldsymbol{p}}$.

The eigenvalue σ_m may also be characterized by the following *minimum-maximum principle*:

The eigenvalue σ_m is the smallest value that can be assumed by the maximum of the quadratic form (3.3–1) *for a unit vector which also satisfies $m-1$ arbitrary conditions of the form*

$$eq_1 = \ldots = eq_{m-1} = 0. \tag{3.3–11}$$

We take a set of vectors $q_1, \ldots, q_{m-1}$. In combination with $p_{m+1}, \ldots, p_n$ they span a linear space $\mathfrak{R}_k$ of dimension $k < n$. As a consequence there is a vector $y \neq o$ not included in $\mathfrak{R}_k$. If P denotes the projection operator associated with $\mathfrak{R}_k$ (section 2.2.1) the vector $x = y - \mathsf{P}y$ is not the zero vector and is orthogonal to $\mathfrak{R}_k$. Hence we may suppose that x is a unit vector e. We apply Euler's theorem to this vector, observing that $\theta_{m+1} = \ldots = \theta_n = \frac{1}{2}\pi$. Hence

$$\begin{aligned} \sigma = e\mathsf{A}e &= \sigma_1 \cos^2 \theta_1 + \ldots + \sigma_m \cos^2 \theta_m \geqq \sigma_m(\cos^2 \theta_1 + \ldots + \cos^2 \theta_m) \\ &= \sigma_m \sum_{h=1}^{n} \cos^2 \theta_h = \sigma_m. \end{aligned}$$

Thus we see that for certain vectors satisfying (3.3–11) the value of the quadratic form (3.3–1) is not less than σ_m and *a fortiori* is the maximum value for all vectors satisfying (3.3–11) is not less than σ_m. On the other hand, since the vectors $q_1, \ldots, q_{m-1}$ are arbitrary and we may replace them by $p_1, \ldots, p_{m-1}$, we infer that p_m is a vector satisfying the conditions stated and σ_m occurs among the values for vectors satisfying (3.3–11).

3.4 – The principal angles between two finite dimensional spaces

3.4.1 – The projection operator

It is our aim to illustrate the theory of a symmetric operator in a geometric problem. To this end we shall consider in more detail some properties of the projection operator introduced in section 2.2.1.

Let $\mathfrak{R}_m$ denote an m-dimensional subspace of a metric space $\mathfrak{R}$. With $\mathfrak{R}_m$ is associated an operator P such that every vector x of $\mathfrak{R}$ can be decomposed into two vectors $\mathsf{P}x$ and $x - \mathsf{P}x$, where $\mathsf{P}x$ is included in $\mathfrak{R}_m$ and $x - \mathsf{P}x$ is orthogonal to $\mathfrak{R}_m$. First we shall prove:

The operator P is a linear operator.

Let u and v denote two arbitrary vectors of $\mathfrak{R}$ and a and b two arbitrary numbers. From the decompositions

$$u = \mathsf{P}u + (u - \mathsf{P}u),$$
$$v = \mathsf{P}v + (v - \mathsf{P}v),$$

follows that

$$ua + vb = (\mathsf{P}u)a + (\mathsf{P}v)b + (u - \mathsf{P}u)a + (v - \mathsf{P}v)b.$$

Since $(u - \mathsf{P}u)a + (v - \mathsf{P}v)b$ is also orthogonal to $\mathfrak{R}_m$ and, moreover, the decomposition is unique, we may conclude that

$$\mathsf{P}(ua + vb) = (\mathsf{P}u)a + (\mathsf{P}v)b.$$

It is clear that the vectors orthogonal to $\mathfrak{R}_m$ are singular for P. Hence the operator P is then and only then non-singular when $\mathfrak{R}$ and $\mathfrak{R}_m$ coincide.

Since $\mathsf{P}x$ is in $\mathfrak{R}_m$ we evidently have

$$\mathsf{P}^2 x = \mathsf{P}x. \tag{3.4–1}$$

This equation expresses the fact that P is *idempotent*, that is to say, the iterated operator coincides with the operator itself.

Next we shall prove that P *is symmetric*. Let x and u denote arbitrary vectors. Since $\mathsf{P}y$ is in $\mathfrak{R}_m$ and $x - \mathsf{P}x$ is orthogonal to $\mathfrak{R}_m$ we have

$$(x - \mathsf{P}x)\mathsf{P}y = o.$$

It follows that

$$x\mathsf{P}y = \mathsf{P}x\mathsf{P}y = \mathsf{P}y\mathsf{P}x = y\mathsf{P}x.$$

The theory of the eigenvalues of the projection operator is not very exciting. In fact, the eigenvectors of P are the vectors of $\mathfrak{R}_m$ and the vectors orthogonal to $\mathfrak{R}_m$. They correspond to the eigenvalues 1 and 0 respectively.

3.4.2 – The principal angles

The angle between a vector x and a finite dimensional subspace $\mathfrak{R}_m$ of the metric space $\mathfrak{R}$ may be defined as the angle between this vector and its projection on the subspace. When the vector is orthogonal to $\mathfrak{R}_m$ then the projection is the zero vector and we define the angle as $\frac{1}{2}\pi$. A general theorem states:

The angle θ between a vector $x \neq o$ and a finite dimensional subspace does not exceed $\frac{1}{2}\pi$. It is determined by

$$x\mathsf{P}x = |x|^2 \cos^2 \theta, \tag{3.4–2}$$

where $\mathsf{P}x$ denotes the projection of x on the subspace.

The case $\mathsf{P}x = o$ has already been considered. Under the assumption

$\mathsf{P}x \neq o$ we have, by virtue of (3.4–1), taking account of the symmetry of P,

$$x\mathsf{P}x = x\mathsf{P}(\mathsf{P}x) = \mathsf{P}x\mathsf{P}x = |\mathsf{P}x|^2 > 0 \tag{3.4–3}$$

and since also

$$x\mathsf{P}x = |x||\mathsf{P}x| \cos\theta \tag{3.4–4}$$

it follows that $\cos\theta > 0$, i.e., $\theta < \frac{1}{2}\pi$.

Equating the right-hand numbers of (3.4–3) and (3.4–4) we get on dividing by $|\mathsf{P}x|$:

$$|\mathsf{P}x| = |x| \cos\theta. \tag{3.4–5}$$

Inserting this into the right-hand member of (3.4–4) the desired result (3.4–2) is obtained.

Suppose now that $\mathfrak{R}_m$ and $\mathfrak{R}_n$ are two finite dimensional subspaces of a metric space $\mathfrak{R}$, the dimensions being indicated by the subscripts. The projection operator associated with $\mathfrak{R}_m$ will be denoted by P, the projection operator associated with $\mathfrak{R}_n$ by Q. Let x be a vector of $\mathfrak{R}_m$. Then $\mathsf{Q}x$ is a vector of $\mathfrak{R}_n$ and $\mathsf{PQ}x$ is again a vector of $\mathfrak{R}_m$.

The operator PQ *is a symmetric linear operator defined throughout* $\mathfrak{R}_m$.

The linearity is evident. From the symmetry of P and Q follows, when x and y are in $\mathfrak{R}_m$:

$$x\mathsf{PQ}y = x\mathsf{P}(\mathsf{Q}y) = \mathsf{Q}y\mathsf{P}x = \mathsf{P}x\mathsf{Q}y = x\mathsf{Q}y = y\mathsf{Q}x = y\mathsf{PQ}x.$$

Similarly QP denotes an operator which is defined throughout $\mathfrak{R}_n$. Now we shall prove the remarkable theorem:

The operators PQ *and* QP *have the same rank.*

We refer to section 3.2.3 where we proved the theorem stating that the rank r of a symmetric linear operator in a finite dimensional metric vector space is the maximal number of linearly independent eigenvectors belonging to non-vanishing eigenvalues. Let $\sigma \neq 0$ denote an eigenvalue of PQ. Then there exists a vector $x \neq o$ such that

$$\mathsf{PQ}x = \sigma x.$$

Hence

$$\mathsf{QPQ}x = \sigma \mathsf{Q}x. \tag{3.4–6}$$

This expresses the fact that $\mathsf{Q}x$ is an eigenvector of QP, for $\mathsf{Q}x = o$ would imply $\mathsf{PQ}x = o$. Further let $\underset{1}{p}, \ldots, \underset{r}{p}$ be a system of mutually orthogonal principal directions of PQ belonging to the non-zero eigenvalues $\underset{1}{\sigma}, \ldots, \underset{r}{\sigma}$. Then $\mathsf{Q}\underset{1}{p}, \ldots, \mathsf{Q}\underset{r}{p}$ are eigenvectors of QP belonging to the same eigenvalues. Suppose that a relation

$$(\mathsf{Q}\underset{1}{p})\underset{1}{\xi} + \ldots + (\mathsf{Q}\underset{r}{p})\underset{r}{\xi} = 0$$

holds. On applying the operator P it follows that also

$$(\mathsf{PQ}\underset{1}{\boldsymbol{p}})\underset{1}{\xi}+\dots+(\mathsf{PQ}\underset{r}{\boldsymbol{p}})\underset{r}{\xi}=\boldsymbol{o}$$

or

$$\underset{1}{\boldsymbol{p}}\underset{1}{\sigma}\underset{1}{\xi}+\dots+\underset{r}{\boldsymbol{p}}\underset{r}{\sigma}\underset{r}{\xi}=0,$$

Since $\underset{1}{\boldsymbol{p}},\dots,\underset{r}{\boldsymbol{p}}$ are linearly independent and none of the numbers $\underset{1}{\sigma},\dots,\underset{r}{\sigma}$ vanish, we infer that $\underset{1}{\xi}=\dots=\underset{r}{\xi}=0$. Hence $\mathsf{Q}\underset{1}{\boldsymbol{p}},\dots,\mathsf{Q}\underset{r}{\boldsymbol{p}}$ are linearly independent and are eigenvectors of the operator QP belonging to non-vanishing eigenvalues. As a consequence the rank of QP is not less than the rank of PQ. But we may interchange the roles of P and Q and then the assertion follows at once.

As a by-product we find that this common rank satisfies the inequality

$$r\leqq\min(m,n). \qquad (3.4\text{–}7)$$

We are now sufficiently prepared to measure the inclination between the spaces $\mathfrak{R}_m$ and $\mathfrak{R}_n$. Let $\boldsymbol{x}\neq\boldsymbol{o}$ denote an eigenvector of PQ belonging to an eigenvalue $\sigma\neq0$. Then also $\mathsf{Q}\boldsymbol{x}\neq\boldsymbol{o}$ and in view of the last theorem of the previous section the angle ω between $\boldsymbol{x}$ and $\mathsf{Q}\boldsymbol{x}$ is an acute angle determined by

$$\boldsymbol{x}\mathsf{Q}\boldsymbol{x}=|\boldsymbol{x}|^2\cos^2\omega. \qquad (3.4\text{–}8)$$

Since $\boldsymbol{x}$ is included in $\mathfrak{R}_m$ we also have

$$\boldsymbol{x}\mathsf{Q}\boldsymbol{x}=\mathsf{P}\boldsymbol{x}\mathsf{Q}\boldsymbol{x}=\mathsf{Q}\boldsymbol{x}\mathsf{P}\boldsymbol{x}=\boldsymbol{x}\mathsf{PQ}\boldsymbol{x}=\sigma|\boldsymbol{x}|^2.$$

Hence

$$\sigma=\cos^2\omega. \qquad (3.4\text{–}9)$$

An angle of this kind will be referred to as a *principal angle* between $\mathfrak{R}_m$ and $\mathfrak{R}_n$. Now we observe that in $\mathfrak{R}_m$ we can find an orthonormal system of eigenvectors $\underset{1}{\boldsymbol{p}},\dots,\underset{r}{\boldsymbol{p}}$ corresponding to r non-vanishing eigenvalues $\underset{1}{\sigma},\dots,\underset{r}{\sigma}$ of the operator PQ. The corresponding principal angles are determined by

$$\cos^2\underset{h}{\omega}=\underset{h}{\sigma}, \qquad h=1,\dots,r. \qquad (3.4\text{–}10)$$

Some of these angles may have the value zero; they correspond to eigenvalues equal to unity. In that case $\mathfrak{R}_m$ and $\mathfrak{R}_n$ have non-zero vectors in common.

It may happen that PQ has more eigenvalues than those listed in (3.4–10). They are necessarily equal to zero and the eigenvectors belonging to them are orthogonal to $\mathfrak{R}_n$. In the particular case that the rank of PQ is zero all vectors of $\mathfrak{R}_m$ are orthogonal to $\mathfrak{R}_n$. Then we say that $\mathfrak{R}_m$ and $\mathfrak{R}_n$ are *absolutely orthogonal.* Summing up we may state:

Two finite dimensional subspaces $\mathfrak{R}_m$ *and* $\mathfrak{R}_n$ *in a metric vector space* $\mathfrak{R}$ *determine* $r = \min(m, n)$ *principal angles* $\underset{1}{\omega}, \ldots, \underset{r}{\omega}$. *The first space contains* $m-r$ *linearly independent vectors orthogonal to* $\mathfrak{R}_n$ *and the second contains* $n-r$ *vectors orthogonal to* $\mathfrak{R}_m$ *or in other words: the first space includes a subspace* $\mathfrak{R}_{m-r}$ *absolutely orthogonal to* $\mathfrak{R}_n$ *and the second includes a subspace* $\mathfrak{R}_{n-r}$ *absolutely orthogonal to* $\mathfrak{R}_m$.

Finally we mention: If a vector x in $\mathfrak{R}_m$ makes the angles $\underset{h}{\vartheta}$, $h = 1, \ldots, r$, with the selected eigenvectors $\underset{h}{p}$ and the angle ω with its projection on $\mathfrak{R}_n$, then

$$\cos^2 \omega = \sum_h \cos^2 \underset{h}{\omega} \cos^2 \underset{h}{\vartheta}, \qquad h = 1, \ldots, r. \tag{3.4–11}$$

This is a direct consequence of Euler's theorem.

Extremal properties of the angle ω can be stated on applying the results of section 3.3.3.

3.4.3 – Orthogonal operators

The foregoing results may be utilized for the description of some properties of so-called orthogonal operators. An operator A in a metric vector space is said to be orthogonal if

$$\mathsf{A}x\mathsf{A}y = xy, \tag{3.4–12}$$

where x and y are arbitrary vectors. First we state:

The only eigenvalues of an orthogonal operator are the numbers ± 1.

For, from

$$\mathsf{A}x = \sigma x$$

follows

$$\mathsf{A}x\mathsf{A}x = \sigma^2 xx,$$

whence, by virtue of (3.3–12), $\sigma^2 = 1$, i.e., $\sigma = \pm 1$.

The second theorem of section 3.2.1 also holds for orthogonal operators, that is to say:

An orthogonal operator sends every vector orthogonal to an eigenvector into a vector which is again orthogonal to this eigenvector.

Let

$$\mathsf{A}x = \pm x$$

and $yx = 0$. By (3.3–12) $\mathsf{A}y\mathsf{A}x = 0$ also or $\pm(\mathsf{A}y)x = 0$. This proves the assertion.

Henceforth we suppose that the space is finite dimensional of dimension n. According to the first theorem of this section there are no non-zero singular vectors of the operator **A** and its rank is therefore equal to n. Suppose that there are m eigenvalues. This means that the characteristic equation

$$\det [a_{\lambda\mu} - \sigma g_{\lambda\mu}] = 0 \tag{3.4–13}$$

possesses m real roots, these being ± 1, and consequently $n-m$ conjugate complex roots, for the coefficients of the equation are real. Hence $n-m$ is an even number $2k$. Next we consider the $2k$-dimensional space orthogonal to the eigenspace of dimension m. This space is transformed into itself. A k-dimensional subspace of it is transformed into another k-dimensional subspace. According to the results of the previous section the product of the projection operators associated with these subspaces possesses k eigenvalues $\underset{h}{\sigma}$, $h = 1, \ldots, \mathrm{k}$. The angles determined by

$$\cos^2 \underset{h}{\vartheta} = \underset{h}{\sigma}, \qquad h = 1, \ldots, \mathrm{k}, \tag{3.4–14}$$

are called the angles of rotation of the operator **A**. Summing up we may state:

An orthogonal operator in an n-dimensional metric vector space possesses an m-dimensional invariant space and determines $\frac{1}{2}(n-m)$ angles of rotation, where m denotes the number of real roots of the characteristic equation associated with the operator.

CHAPTER 4

TENSORS

The bilinear forms studied in the previous chapter are particular cases of the so-called multilinear forms which are real-valued functions of several vector variables restricted to the conditions of linearity. It is common to denote these mathematical entities by the name "tensor", because they appear, for instance, in the theory of elasticity and measure the deformations of an elastic body under the influence of external forces. In finite dimensional metric vector spaces they are characterized by a system of components which are transformed in a particular way when we change the frame of reference. Tensors can also be combined by means of simple algebraic operations and give rise to a body of rules, the tensor algebra. The tensors are used to state geometrical facts in a concise way.

4.1 – Multilinear forms

4.1.1 – DEFINITION

By a *multilinear form* is understood a function of a certain number of vector variables such that the function is linear in each variable separately. This means, when one of the variables is replaced by the vector $\boldsymbol{u}a+\boldsymbol{v}b$, then the value of the form is the value for the vector $\boldsymbol{u}$ times a plus the value for the vector $\boldsymbol{v}$ times b. It is clear that a bilinear form is a multilinear form in two vector variables. The number of variables occurring in the form is called its *valency*. We do not want to exclude forms of zero valency. They may be identified with constants.

We have adopted the name "multilinear form" in order to exhibit the particular functional character. It is, however, more common to denote a multilinear form by the name *tensor*. In subsequent chapters we shall use this name exclusively.

Suppose now that the space is a finite dimensional vector space and let (κ) denote a frame. The values of the bilinear form which are taken when the variables are replaced by elements of the frame are called the *components of the tensor* with respect to the frame. The designation of the components will be illustrated in the example of a trilinear form

$$A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}). \tag{4.1–1}$$

Replacing $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ by $\boldsymbol{x}_\lambda, \boldsymbol{x}_\mu, \boldsymbol{x}_\nu$ respectively we obtain a number denoted by

$$a_{\lambda\mu\nu} = A(\boldsymbol{x}_\lambda, \boldsymbol{x}_\mu, \boldsymbol{x}_\nu), \tag{4.1–2}$$

where λ, μ, ν run through the numbers $1, \ldots, n$. The numbers (4.1–2) are called the *purely covariant components* of the tensor. The *purely contravariant components* are obtained by replacing the variables by $\boldsymbol{x}^\lambda$, $\boldsymbol{x}^\mu$, $\boldsymbol{x}^\nu$, that is to say

$$a^{\lambda\mu\nu} = A(\boldsymbol{x}^\lambda, \boldsymbol{x}^\mu, \boldsymbol{x}^\nu). \tag{4.1–3}$$

In addition we have a variety of types of *mixed components*, e.g.,

$$a_\lambda{}^{\mu\nu} = A(\boldsymbol{x}_\lambda, \boldsymbol{x}^\mu, x^\nu), \qquad a^\lambda{}_\mu{}^\nu = A(\boldsymbol{x}^\lambda, \boldsymbol{x}_\mu, \boldsymbol{x}^\nu), \text{ etc.} \tag{4.1–4}$$

Thus we see that the components are denoted by a symbol provided with a number of upper and lower indices equal to the valency of the multilinear form.

The value of the trilinear form (4.1–1) for the vectors $\boldsymbol{u} = \boldsymbol{x}_\kappa u^\kappa$, $\boldsymbol{v} = \boldsymbol{x}_\kappa v^\kappa$, $\boldsymbol{w} = \boldsymbol{x}_\kappa w^\kappa$ may be expressed in terms of the components as

$$A(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = a_{\lambda\mu\nu} u^\lambda v^\mu w^\nu.$$

If we should have taken $\boldsymbol{u} = \boldsymbol{x}^\kappa u_\kappa$, $\boldsymbol{v} = \boldsymbol{x}_\kappa v^\kappa$, $\boldsymbol{w} = \boldsymbol{x}^\kappa w_\kappa$ the result would be

$$A(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = a^\lambda{}_\mu{}^\nu u_\lambda v^\mu w_\nu$$

and so on. A set of components of a certain type determines all other types. In fact, in view of (2.2–6) and (2.2–8) we have, for instance,

$$\begin{aligned} a^\lambda{}_{\mu\nu} &= A(\boldsymbol{x}^\lambda, \boldsymbol{x}_\mu, \boldsymbol{x}_\nu) = A(g^{\lambda\kappa}\boldsymbol{x}_\kappa, \boldsymbol{x}_\mu, \boldsymbol{x}_\nu) \\ &= g^{\lambda\kappa} A(\boldsymbol{x}_\kappa, \boldsymbol{x}_\mu, \boldsymbol{x}_\nu) = g^{\lambda\kappa} a_{\kappa\mu\nu}. \end{aligned}$$

Conversely

$$a_{\lambda\mu\nu} = g_{\lambda\kappa} a^\kappa{}_{\mu\nu}.$$

This is the process of raising and lowering the indices.

4.1.2 – The law of transformation

Next we introduce a second frame (κ'). Taking account of (1.1–11) and (2.2–22) we easily deduce equations of transformation like

$$\begin{aligned} a_{\lambda'\mu'\nu'} &= a_{\lambda\mu\nu} q^\lambda_{\lambda'} q^\mu_{\mu'} q^\nu_{\nu'}, \\ a_{\lambda'}{}^{\mu'\nu'} &= a_\lambda{}^{\mu\nu} q^\lambda_{\lambda'} q^{\mu'}_\mu q^{\nu'}_\nu, \end{aligned}$$

etc. In the case of a tensor of valency $h+k$ we may write

$$\boxed{a^{\lambda'_1 \ldots \lambda'_h}{}_{\mu'_1 \ldots \mu'_k} = a^{\lambda_1 \ldots \lambda_h}{}_{\mu_1 \ldots \mu_k} q^{\lambda'_1}_{\lambda_1} \cdots q^{\lambda'_h}_{\lambda_h} q^{\mu_1}_{\mu'_1} \cdots q^{\mu_k}_{\mu'_k}.} \tag{4.1–5}$$

This is only one type of components, but, as we already pointed out in the example of a trilinear form, many possible arrangements of the indices present themselves. In all these cases, however, the equations of transformation are essentially the same.

We started with multilinear forms and arrived at the components and their laws of transformation. In most applications, however, the components are given. Suppose that to every frame (κ) we let correspond a system of n^{h+k} numbers $a^{\lambda_1\cdots\lambda_h}{}_{\mu_1\cdots\mu_k}$, such that between the numbers corresponding to the frames (κ) and (κ') the relations (4.1–4) hold. We assert that the numbers corresponding to any frame (κ) are the components of a tensor with respect to this frame.

In order to prove this assertion we consider the form

$$a^{\lambda_1\cdots\lambda_h}{}_{\mu_1\cdots\mu_k}\underset{1}{x}_{\lambda_1}\cdots\underset{h}{x}_{\lambda_h}\underset{h+1}{x}^{\mu_1}\cdots\underset{h+k}{x}^{\mu_k} \tag{4.1-6}$$

where $\underset{1}{x}_{\lambda_1},\ldots,\underset{h}{x}_{\lambda_h}$ are the covariant components of the vectors $\underset{1}{\mathbf{x}},\ldots,\underset{h}{\mathbf{x}}$ and $\underset{h+1}{x}^{\mu_1},\ldots,\underset{h+k}{x}^{\mu_k}$ are the contravariant components of the vectors $\underset{h+1}{\mathbf{x}}.\ldots,\underset{h+k}{\mathbf{x}}$.

On account of (4.1–5), (1.1–18) and (2.2–23) we have:

$$\begin{aligned}
&a^{\lambda'_1\cdots\lambda'_h}{}_{\mu'_1\cdots\mu'_k}\underset{1}{x}_{\lambda'_1}\cdots\underset{h}{x}_{\lambda'_h}\underset{h+1}{x}^{\mu'_1}\cdots\underset{h+1}{x}^{\mu'_k}\\
&= a^{\lambda_1\cdots\lambda_h}{}_{\mu_1\cdots\mu_k}q^{\lambda'_1}_{\lambda_1}\cdots q^{\lambda'_h}_{\lambda_h}q^{\mu_1}_{\mu'_1}\cdots q^{\mu_k}_{\mu'_k}\underset{1}{x}_{\lambda'_1}\cdots\underset{h}{x}_{\lambda'_h}\underset{h+1}{x}^{\mu'_1}\cdots\underset{h+k}{x}^{\mu'_k}\\
&= a^{\lambda_1\cdots\lambda_h}{}_{\mu_1\cdots\mu_k}\underset{1}{x}_{\lambda_1}\cdots\underset{h}{x}_{\lambda_h}\underset{h+1}{x}^{\mu_1}\cdots\underset{h+k}{x}^{\mu_k}.
\end{aligned}$$

Hence the values of the expressions (4.1–5) evaluated in various frames are the same and may therefore be denoted by

$$A(\underset{1}{\mathbf{x}},\ldots,\underset{h}{\mathbf{x}},\underset{h+1}{\mathbf{x}},\ldots,\underset{h+k}{\mathbf{x}}), \tag{4.1.7}$$

being a function of the vector variables $\underset{1}{\mathbf{x}},\ldots,\underset{h}{\mathbf{x}},\underset{h+1}{\mathbf{x}},\ldots,\underset{h+k}{\mathbf{x}}$. It is immediately clear, that this function is a multilinear form.

A particular case is that of a tensor of valency one. It is characterized by a set of n components a^λ or a_λ obeying the rules of transformation (2.2–23) and (1.1–18). Hence *a tensor of valency one may be identified with a vector.*

A tensor of valency zero is given when to every frame (κ) corresponds a number $\overset{(\kappa)}{\varphi}$, such that

$$\overset{(\kappa')}{\varphi} = \overset{(\kappa)}{\varphi}. \tag{4.1-8}$$

A tensor of this kind is called a *scalar.*

A tensor of utmost importance is that which is represented by the inner product $\boldsymbol{xy}$ of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. Its components are $g_{\lambda\mu}$, $g^{\lambda\mu}$ and δ^{λ}_{μ}. This tensor will be referred to as the *metric tensor*.

4.2 – Tensor algebra

4.2.1 – Addition and multiplication

There are some simple algebraic rules for combining given tensors with others. We shall illustrate them with special examples, the general case being omitted, but the pattern is the same as for our simple examples.

First we define the *sum* $C(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ of two tensors of the same valency (the only case where the addition has a meaning)

$$C(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) + B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \tag{4.2–1}$$

as the sum of the functions $A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ and $B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$. Introducing components we have, for instance,

$$c^{\kappa}{}_{\lambda\mu} = a^{\kappa}{}_{\lambda\mu} + b^{\kappa}{}_{\lambda\mu} \tag{4.2–2}$$

and similar results in the case of another arrangement of the indices.

Next we consider *the tensor product* (or briefly *the product*) of two tensors, the valency of which need not be the same, e.g., $A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$, $B(\boldsymbol{u}, \boldsymbol{v})$ as the product

$$C(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{v}) = A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) B(\boldsymbol{u}, \boldsymbol{v}) \tag{4.2–3}$$

of these functions. Notice that the variables appearing in the different factors are also different. On inserting basis vectors we have, for instance,

$$C(\boldsymbol{x}_{\lambda}, \boldsymbol{x}^{\mu}, \boldsymbol{x}_{\nu}, \boldsymbol{x}_{\kappa}, \boldsymbol{x}^{\rho}) = A(\boldsymbol{x}_{\lambda}, \boldsymbol{x}^{\mu}, \boldsymbol{x}_{\nu}) B(\boldsymbol{x}_{\kappa}, \boldsymbol{x}^{\rho}),$$

that is to say, the numbers

$$c_{\lambda}{}^{\mu}{}_{\nu\kappa}{}^{\rho} = a_{\lambda}{}^{\mu}{}_{\nu} b_{\kappa}{}^{\rho} \tag{4.2–4}$$

are components of the product of the tensors represented by the components $a_{\lambda}{}^{\mu}{}_{\nu}$ and $b_{\kappa}{}^{\rho}$.

Addition and multiplication are so-called *covariant operations*. This means that these processes applied to components in a second frame (κ') yield components which are obtained by direct transformation. Thus, for instance,

$$\begin{aligned} c^{\kappa'}{}_{\lambda'\mu'} &= a^{\kappa'}{}_{\lambda'\mu'} + b^{\kappa'}{}_{\lambda'\mu'} = a^{\kappa}{}_{\lambda\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} + b^{\kappa}{}_{\lambda\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} \\ &= (a^{\kappa}{}_{\lambda\mu} + b^{\kappa}{}_{\lambda\mu}) q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'} = c^{\kappa}{}_{\lambda\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'}. \end{aligned}$$

A tensor, whose components with respect to a given frame (κ) are zero, has zero components with respect to every frame. This is the reason why tensors are important entities, for their vanishing expresses a fact independent of a particular frame.

4.2.2 – Contraction and transvection

Consider a tensor represented by the components

$$a^{\kappa}{}_{\lambda\mu}. \tag{4.2–5}$$

We perform on it an operation, called *contraction*, consisting of equating an upper and a lower index and summing on these identified indices. Thus contraction on the indices κ and λ means the formation of the set of numbers

$$b_{\mu} = a^{\kappa}{}_{\kappa\mu}. \tag{4.2–6}$$

We shall prove that these numbers are again the components of a tensor. Performing the contraction on the components, with respect to a frame (κ') we find

$$\begin{aligned} b_{\kappa'} &= a^{\kappa'}{}_{\kappa'\mu'} = a^{\kappa}{}_{\lambda\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\kappa'} q^{\mu}_{\mu'} \\ &= a^{\kappa}{}_{\lambda\mu} \delta^{\lambda}_{\kappa} q^{\mu}_{\mu'} = a^{\kappa}{}_{\kappa\mu} q^{\mu}_{\mu'} = b_{\mu} q^{\mu}_{\mu'}, \end{aligned}$$

where we have made use of (1.1–17).

We have illustrated the process in a particular example, but the general pattern is the same.

Contraction can only be applied to the mixed components of a tensor As we pointed out it is a covariant process. Applied to the components of a tensor product on indices occurring in different factors it is called *transvection*. We have already encountered this process in many examples, in particular in the cases of lowering or raising an index.

Transvection applied to the tensor product of two vectors represented by their components a_{λ}, b^{μ}, this product being the tensor $a_{\lambda} b^{\mu}$, yields the number $a_{\lambda} b^{\lambda}$, the inner product of the vectors. Transvection applied to a tensor and the Kronecker delta yields the re-naming of an index:

$$a_{\lambda}{}^{\mu} \delta^{\lambda}_{\kappa} = a_{\kappa}{}^{\mu},$$

a process which has also been applied several times.

4.2.3 – The quotient rule

Suppose for every frame (κ) we are given a system of n^3 numbers $a^{\kappa}{}_{\lambda\mu}$. In order to prove that they are the components of a tensor we have to verify that between the system corresponding to (κ) and that corresponding to (κ') a relation of the type (4.1–4) holds. In many cases, however, this may be rather cumbersome, but often the work may be reduced and facilitated by applying the following device: Let v^{μ} denote the contravariant components of an arbitrary vector. If the numbers

$$a^{\kappa}{}_{\lambda\mu} v^{\mu} \tag{4.2–7}$$

turn out to be the components of a tensor then the numbers $a^{\kappa}{}_{\lambda\mu}$ are also the components of a tensor.

The proof is easy. By hypothesis we have

$$a^{\kappa'}{}_{\lambda'\mu'} v^{\mu'} = a^{\kappa}{}_{\lambda\mu} v^{\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'}.$$

Since v^{μ} are the components of a vector we also have

$$a^{\kappa'}{}_{\lambda'\mu'} v^{\mu} q^{\mu'}_{\mu} = a^{\kappa}{}_{\lambda\mu} v^{\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'}.$$

These equations hold for all values of v^{μ}. As a consequence we have

$$a^{\kappa'}{}_{\lambda'\mu'} q^{\mu'}_{\mu} = a^{\kappa}{}_{\lambda\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'},$$

or

$$a^{\kappa'}{}_{\lambda'\mu'} = a^{\kappa}{}_{\lambda\mu} q^{\kappa'}_{\kappa} q^{\lambda}_{\lambda'} q^{\mu}_{\mu'}$$

and this completes the proof.

Since the components $a^{\kappa}{}_{\lambda\mu}$ have been obtained from the expressions (4.2–7) by formal division by v^{μ} the theorem stated above is often called *the quotient rule.*

As a simple example we consider the Kronecker delta. When v^{μ} is an arbitrary vector then also

$$v^{\lambda} = \delta^{\lambda}_{\mu} v^{\mu}$$

Hence the δ^{λ}_{μ} are the components of a tensor. In this case, however, a direct verification is also very easy.

4.2.4 – Vectorial tensors

As we pointed out in section 4.1.2 a tensor may be given by its components with respect to a frame (κ). In geometric applications we also encounter similar entities, but then the components are vectors. Stated in more detail: Suppose that to every frame (κ) corresponds a set of n^h vectors

$$\boldsymbol{x}_{\lambda_1 \ldots \lambda_h} \tag{4.2–8}$$

such that between this system and the system corresponding to a frame (κ') exist the relations

$$\boldsymbol{x}_{\lambda'_1 \ldots \lambda'_h} = \boldsymbol{x}_{\lambda_1 \ldots \lambda_h} q^{\lambda_1}_{\lambda'_1} \cdots q^{\lambda_h}_{\lambda'_h}, \tag{4.2–9}$$

then we shall say that the vectors (4.2–8) are the components with respect to (κ) of a *vectorial tensor of valency h.*

We restrict ourselves to vectorial tensors provided with covariant indices only, for other types do not occur, except the vectorial tensor $\boldsymbol{x}^{\lambda}$, the components of which are the elements of the reciprocal basis of the basis $\boldsymbol{x}_{\lambda}$.

The algebraic rules for tensors also hold for vectorial tensors, it being understood, that the product of two vectorial tensors provides an ordinary tensor. Thus, for instance, the tensor product of a basis $\boldsymbol{x}_\lambda$ by itself yields the metric tensor $g_{\lambda\mu} = \boldsymbol{x}_\lambda \boldsymbol{x}_\mu$.

The components (4.2–8) may be completed to a multilinear form

$$\boldsymbol{x}_{\lambda_1 \ldots \lambda_h} \underset{1}{x}^{\lambda_1} \ldots \underset{h}{x}^{\lambda_h}, \tag{4.2–10}$$

the value now being a vector.

4.3 – Multivectors

4.3.1 – Symmetry and antisymmetry

A tensor is called *symmetric* with respect to a pair of its vector variables when it remains unaltered by interchanging these variables. Thus, for instance, when

$$A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = A(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{z}) \tag{4.3–1}$$

identically, then this tensor is symmetric with respect to the pair $\boldsymbol{x}, \boldsymbol{y}$. Applied to the components we find

$$A(\boldsymbol{x}_\lambda, \boldsymbol{x}_\mu, \boldsymbol{x}_\nu) = A(\boldsymbol{x}_\mu, \boldsymbol{x}_\lambda, \boldsymbol{x}_\nu), \tag{4.3–2}$$

that is to say

$$a_{\lambda\mu\nu} = a_{\mu\lambda\nu}. \tag{4.3–3}$$

This can also be expressed by saying that the tensor is symmetric with respect to its indices λ and μ.

If the tensor changes into its negative when two of its variables are interchanged then we say, that the tensor is *antisymmetric* or *alternating* with respect to these variables. Thus for instance, if

$$A(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = -A(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{z}) \tag{4.3–4}$$

identically, that is to say, if

$$a_{\lambda\mu\nu} = -a_{\mu\lambda\nu}, \tag{4.3–5}$$

then this tensor is antisymmetric with respect to the pair $\boldsymbol{x}, \boldsymbol{y}$ or, with respect to the indices λ and μ.

It should be noticed that it is sufficient to verify (4.3–2) or (4.3–5) in one frame (κ). More generally, if a relation

$$\xi a_{\lambda\mu\nu} + \eta a_{\mu\lambda\nu} = 0 \tag{4.3–6}$$

holds in one frame, then it holds in every frame. This follows since in a system (κ) we have

$$\begin{aligned} \xi a_{\lambda'\mu'\nu'} + \eta a_{\mu'\lambda'\nu'} &= \xi a_{\lambda\mu\nu} q^\lambda_{\lambda'} q^\mu_{\mu'} q^\nu_{\nu'} + \eta a_{\mu\lambda\nu} q^\mu_{\mu'} q^\lambda_{\lambda'} q^\nu_{\nu'} \\ &= (\xi a_{\lambda\mu\nu} + \eta a_{\mu\lambda\nu}) q^\lambda_{\lambda'} q^\mu_{\mu'} q^\nu_{\nu'} = 0. \end{aligned}$$

In many cases *cyclic relations* of the type

$$a_{\lambda\mu\nu} + a_{\mu\nu\lambda} + a_{\nu\lambda\mu} = 0 \tag{4.3-7}$$

occur. It is at once clear, that this condition is covariant, i.e., it holds in every frame if it holds in one frame.

4.3.2 – Bivectors

In many problems a particular type of antisymmetric tensors occur, the *bivectors*. These are antisymmetric tensors of valency two and may be characterized by the conditions

$$f^{\lambda\mu} = -f^{\mu\lambda}. \tag{4.3-8}$$

A special type of bivector is obtained from two vectors a^κ, b^κ, by forming the components

$$f^{\lambda\mu} = \tfrac{1}{2}\begin{vmatrix} a^\lambda & a^\mu \\ b^\lambda & b^\mu \end{vmatrix} = \tfrac{1}{2}(a^\lambda b^\mu - a^\mu b^\lambda). \tag{4.3-9}$$

A bivector of this kind is called *decomposable* or *simple*.

The two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ defining the bivector (4.3–9) generate a *vector plane* consisting of all vectors $\boldsymbol{a}\xi + \boldsymbol{b}\eta$, provided $\boldsymbol{a}$ and $\boldsymbol{b}$ are linearly independent. Suppose that $\overset{*}{\boldsymbol{a}}$ and $\overset{*}{\boldsymbol{b}}$ are included in the plane spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$ and are also linearly independent. Then we have

$$\overset{*}{\boldsymbol{a}} = \boldsymbol{a}\alpha_1 + \boldsymbol{b}\beta_1,$$
$$\overset{*}{\boldsymbol{b}} = \boldsymbol{a}\alpha_2 + \boldsymbol{b}\beta_2$$

and

$$\overset{*}{f}{}^{\lambda\mu} = \tfrac{1}{2}\begin{vmatrix} \overset{*}{a}{}^\lambda & \overset{*}{a}{}^\mu \\ \overset{*}{b}{}^\lambda & \overset{*}{b}{}^\mu \end{vmatrix} = \tfrac{1}{2}\begin{vmatrix} a^\lambda\alpha_1 + b^\lambda\beta_1 & a^\mu\alpha_2 + b^\mu\beta_2 \\ a^\lambda\alpha_1 + b^\lambda\beta_1 & a^\mu\alpha_2 + b^\mu\beta_2 \end{vmatrix}$$

$$= \tfrac{1}{2}\begin{vmatrix} a^\lambda & a^\mu \\ b^\lambda & b^\mu \end{vmatrix}\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}$$

or

$$\overset{*}{f}{}^{\lambda\mu} = f^{\lambda\mu}\begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix}. \tag{4.3-10}$$

Hence, when we replace the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ defining a simple bivector $f^{\lambda\mu}$ by two other vectors $\overset{*}{\boldsymbol{a}}$, $\overset{*}{\boldsymbol{b}}$ in the plane generated by $\boldsymbol{a}$ and $\boldsymbol{b}$ this yields a bivector, the components of which are proportional to the components of the first one.

It is clear that *the components of a simple bivector are zero, if and only if the two vectors defining it are linearly dependent.*

Interesting relations exist between the components of a simple bivector. From

$$\begin{aligned} f^{\kappa\lambda} f^{\mu\nu} &= \tfrac{1}{4}(a^\kappa b^\lambda - a^\lambda b^\kappa)(a^\mu b^\nu - a^\nu b^\mu) \\ &= \tfrac{1}{4}(a^\kappa b^\lambda a^\mu b^\nu - b^\kappa a^\lambda a^\mu b^\nu - a^\kappa b^\lambda b^\mu a^\nu + b^\kappa a^\lambda b^\mu a^\nu) \end{aligned}$$

we find, on performing a cyclic permutation on the indices λ, μ, ν and adding corresponding members of the equations obtained that

$$\boxed{f^{\kappa\lambda} f^{\mu\nu} + f^{\kappa\mu} f^{\nu\lambda} + f^{\kappa\nu} f^{\lambda\mu} = 0.} \tag{4.3–11}$$

Now we may state:

The relations (4.3–11) *express a necessary and sufficient condition for a bivector to be decomposable.*

We only have to establish the sufficiency. We may assume that not all components $f^{\lambda\mu}$ are zero, for in the contrary case the assertion is trivial. Without loss of generality we may suppose that $f^{12} \neq 0$. Next we consider the vectors

$$a^\lambda = f^{1\lambda}, \qquad b^\lambda = f^{2\lambda}, \qquad \lambda = 1, \ldots, n.$$

They generate a bivector with components

$$\overset{*}{f}{}^{\lambda\mu} = \tfrac{1}{2} \begin{vmatrix} a^\lambda & a^\mu \\ b^\lambda & b^\mu \end{vmatrix} = \tfrac{1}{2} \begin{vmatrix} f^{1\lambda} & f^{1\mu} \\ f^{2\lambda} & f^{2\mu} \end{vmatrix} = \tfrac{1}{2}(f^{1\lambda} f^{2\mu} - f^{1\mu} f^{2\lambda}).$$

According to (4.3–11) we have

$$f^{1\lambda} f^{2\mu} + f^{12} f^{\mu\lambda} + f^{1\mu} f^{\lambda 2} = 0.$$

Hence

$$\overset{*}{f}{}^{\lambda\mu} = f^{\lambda\mu} f^{12}.$$

In particular, $\overset{*}{f}{}^{12} = f^{12} f^{12} \neq 0$ and therefore $\boldsymbol{a}$ and $\boldsymbol{b}$ are linearly independent. Thus we see that the numbers $f^{\lambda\mu}$ are proportional to the components of a simple bivector.

4.3.3 – Norm of a bivector

A vector can be pictured as a line segment in a point space (see section 5.1.1) and the norm of the vector is the square of the length of this segment A simple bivector may be pictured as a parallelogram the sides of which are the segments representing the two vectors which define the bivector. Accordingly the norm of the bivector is taken as the square of the area of the parallelogram. When θ is the angle between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ defining the bivector, this area is evidently $|\boldsymbol{a}||\boldsymbol{b}| \sin \theta$.

It is quite natural to take this expression as a definition for the norm of a simple bivector in the abstract case, that is to say, we define the *norm* of the bivector generated by the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by

$$|\boldsymbol{a}|^2|\boldsymbol{b}|^2 \sin^2 \theta = |\boldsymbol{a}|^2||\boldsymbol{b}|^2(1-\cos^2 \theta)| = |\boldsymbol{a}|^2||\boldsymbol{b}|^2-(\boldsymbol{ab})^2. \quad (4.3\text{–}12)$$

In terms of the components with respect to a frame (κ) we find

$$g_{\alpha\lambda}\, a^\alpha a^\lambda g_{\beta\mu}\, b^\beta b^\mu - g_{\alpha\mu}\, a^\alpha b^\mu g_{\beta\lambda}\, b^\beta a^\lambda$$

$$= (g_{\alpha\lambda} g_{\beta\mu} - g_{\alpha\mu} g_{\beta\lambda}) a^\alpha a^\lambda b^\beta b^\mu = \begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} a^\alpha a^\lambda b^\beta b^\mu.$$

Next we observe that the determinant occurring in the last member is alternating in α and β, as well as in λ and μ. Hence we may write

$$\begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} a^\alpha a^\lambda b^\beta b^\mu = \tfrac{1}{2} \begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} (a^\alpha b^\beta - a^\beta b^\alpha) a^\lambda b^\mu$$

$$= \begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} f^{\alpha\beta} a^\lambda b^\mu = \tfrac{1}{2} \begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} f^{\alpha\beta} (a^\lambda b^\mu - a^\mu b^\lambda).$$

Thus we proved:

When $f^{\lambda\mu}$ are the components $\frac{1}{2}(a^\lambda b^\mu - a^\mu b^\lambda)$ of a simple bivector, then

$$\begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} a^\alpha b^\beta a^\lambda b^\mu = \begin{vmatrix} g_{\alpha\lambda} g_{\alpha\mu} \\ g_{\beta\lambda} g_{\beta\mu} \end{vmatrix} f^{\alpha\beta} f^{\lambda\mu}. \quad (4.3\text{–}13)$$

The expression on the right has still a meaning when the $f^{\lambda\mu}$ are the components of a general bivector and can, therefore, be taken as the *norm* of this bivector. A bivector of norm unity is a *unit bivector*. Such a bivector is obtained, for instance, when we take for $\boldsymbol{a}$ and $\boldsymbol{b}$ two orthogonal unit vectors.

4.3.4 – Multivectors

The concept of bivector is a particular case of that of a *multivector* of valency h, also called an *h-vector*, being a tensor of valency h, which is alternating in every pair of adjacent variables. Otherwise stated: An even permutation of the variables leaves the tensor unaltered, an odd permutation changes the sign. In terms of components:

$$f^{\lambda_1 \cdots \lambda_h} = \pm f^{\mu_1 \cdots \mu_h} \quad (4.3\text{–}14)$$

where the plus sign must be used when $\mu_1, \ldots, \mu_h$ is an even permutation of $\lambda_1, \ldots, \lambda_h$ and the minus sign when this permutation is odd.

By a simple or decomposable h-vector is understood a tensor, the components of which are obtained from h vectors $\underset{1}{\boldsymbol{a}}, \ldots, \underset{h}{\boldsymbol{a}}$ on forming the determinants

$$f^{\lambda_1\cdots\lambda_h} = \frac{1}{n!}\begin{vmatrix} \underset{1}{a}^{\lambda_1} & \cdots & \underset{1}{a}^{\lambda_h} \\ \cdot & \cdot & \cdot \\ \underset{h}{a}^{\lambda_1} & \cdots & \underset{h}{a}^{\lambda_h} \end{vmatrix}. \tag{4.3–15}$$

In accordance with (4.3–13) we may define the norm of the multivector as the number

$$\begin{vmatrix} g_{\alpha_1\lambda_1} & \cdots & g_{\alpha_1\lambda_h} \\ \cdot & \cdot & \cdot \\ g_{\alpha_h\lambda_1} & \cdots & g_{\alpha_h\lambda_h} \end{vmatrix} f^{\alpha_1\cdots\alpha_h} f^{\lambda_1\cdots\lambda_h}. \tag{4.3–16}$$

CHAPTER 5

MANIFOLDS IN A METRIC POINT SPACE

Differential geometry is concerned with properties of manifolds in a point space. In the theory the vectors play a dominating part and it is, therefore, natural to base the considerations on the concept of vectors. As a consequence the points take secondary place.

Vectors are usually defined by means of oriented line segments in a point space. The properties of vectors arise from geometrical considerations and they serve to express geometrical theorems in a concise form. In this chapter we shall proceed in a converse direction. The starting-point remains unaltered, that is, we make use of the fact that we are already in possession of the concept of vector. By means of additional axioms we shall introduce the notion of point.

Our main task will be the study of certain subsets in a point space, the manifolds. Although the manifolds are to be defined in terms of points we wish to emphasize the intimate connection between points and vectors. As we shall see, the particular structure of the manifolds finds its roots in the properties of vectors and related concepts.

It is not our intention to consider the most general types of manifolds. Since we are only interested in so-called local properties, a restricted type of manifold is sufficient. The study of manifolds as a whole demands aids provided by modern topological theories.

5.1 – The point space

5.1.1 – RELATION BETWEEN POINTS AND VECTORS

In addition to the vectors of a linear vector space we need a second kind of mathematical entities, hereafter called points, which are related to vectors in a peculiar manner. The points shall be indicated by italic capitals.

It will be convenient to use the notion of "difference" of two points A and B. This only means that in an intuitive geometric model a vector is represented by a line segment, the terminals of which are the given points. Since the segment is provided with an orientation we use the expression "difference", for, like in a difference, the interchange of the points changes the orientation into its opposite. As a matter of fact the difference has at the outset no connection with the arithmetical operation, but in a certain case, however, viz. in the case of a number space to be introduced in section 5.2.2 it has.

The *point space* is described in the following axiomatic way:

Given a vector space $\mathfrak{X}$ there exists a set $\mathscr{X}$ of mathematical entities, satisfying the conditions:

1) To every ordered pair (A, B) of elements of $\mathscr{X}$ corresponds a unique vector of $\mathfrak{X}$, called the *difference* $B-A$ of B and A.

2) $B-A = -(A-B)$.

3) $(C-B)+(B-A) = C-A$.

4) If O is an arbitrary element of $\mathscr{X}$, then to every vector $\boldsymbol{x}$ of $\mathfrak{X}$ corresponds a unique element P in $\mathscr{X}$, such that

$$P-O = \boldsymbol{x}. \tag{5.1–1}$$

A set $\mathscr{X}$ possessing the above properties is called an *affine* point space associated with the vector space $\mathfrak{X}$. The elements of $\mathscr{X}$ are called *points*.

The ordered pair of points (A, B), called an *oriented segment*, determines the vector $B-A$. The point A is the *initial point* and the point B the *terminal point* of the segment. Both points are called *end points*. It is clear that the elementary notion of vectors in a point space is in accordance with our abstract approach.

Two oriented segments are said to be *equipollent* if the corresponding vectors are equal. A vector determines a class of mutually equipollent segments.

Given a vector $\boldsymbol{a}$ we can assign to every point A a point A' such that

$$A'-A = \boldsymbol{a}. \tag{5.1–2}$$

The correspondence $A \to A'$ is named a *translation* of the space X into itself. Thus we may also interpret vectors as translations of a point space.

5.1.2 – Flats

In an affine space $\mathscr{X}$ associated with a linear vector space $\mathfrak{X}$ we consider the set of points Q such that

$$Q-P = \boldsymbol{y}, \tag{5.1–3}$$

where P is a fixed point and $\boldsymbol{y}$ runs through a linear subspace $\mathfrak{Y}$ of $\mathfrak{X}$. A set of this kind will be referred to as a *flat*. Since $P-P = \boldsymbol{o}$ the point P is also included in the flat. As a matter of fact, the whole space $\mathscr{X}$ is a flat. Strictly speaking, a point is also a flat, viz. corresponding to a subspace of $\mathfrak{X}$ consisting of the zero vector only. This trivial case will be excluded in the sequel.

Two flats associated with the same vector space $\mathfrak{Y}$ are called *parallel*; they either coincide or have no point in common. More generally we may say that a flat associated with a vector space $\mathfrak{Y}$ is parallel to a flat associated with a vector space $\mathfrak{Z}$, $\mathfrak{Y}$ and $\mathfrak{Z}$ being subspaces of $\mathfrak{X}$, when either $\mathfrak{Y}$ is included in $\mathfrak{Z}$ or $\mathfrak{Z}$ is included in $\mathfrak{Y}$.

When $\mathfrak{Y}_n$ denotes a finite dimensional vector space of dimension n, we say that any flat associated with $\mathfrak{Y}_n$ is of dimension n.

A *straight line*, briefly a *line*, is a flat associated with a ray; its dimension is unity. An orientation of the ray induces an orientation of the line in the following sense. If one of the classes of the ray is assigned as the positive class, then we say that the point P on the line *precedes* the point Q if and only if the vector $Q-P$ is included in the positive class. Thus a line appears as an ordered set of points.

Two different points A and B determine uniquely a line containing these points, the associated ray being spanned by the vector $B-A$. In fact, all points P of the line are given by

$$P-A = (B-A)q, \tag{5.1–4}$$

where q runs through the set of all real numbers. For $q = 0$ the point P coincides with A, for $q = 1$ the point P is the point B.

A two dimensional flat is usually called a *plane*. Three points A, B and C not on a line are included in just one plane. It is the set of points P given by

$$P-C = (A-C)p+(B-C)q, \tag{5.1–5}$$

where p and q are arbitrary real numbers.

5.1.3 – The coordinates of a point

In a point space $\mathscr{X}$ associated with a vector space $\mathfrak{X}$ we take a point O. By means of the formula (5.1–1) a one-to-one correspondence is established between the set of points $\mathscr{X}$ and the set of vectors $\mathfrak{X}$. This correspondence is a *vectorial coordinate system* with origin O. The vector $\boldsymbol{x}$ is the *coordinate vector of* P with respect to O.

This coordinate vector changes when we take another origin. The equations of transformation are easily found. Let O^* denote a second fixed point and denote the vector O^*-O by $\boldsymbol{a}$. From

$$P-O = (O^*-O)+(P-O^*)$$

it follows at once that

$$\boldsymbol{x} = \boldsymbol{a}+\boldsymbol{x}^*, \tag{5.1–6}$$

where $\boldsymbol{x}^*$ is the coordinate vector of the same point P with respect to the origin O^*.

Next we suppose that the vector space is an n-dimensional space $\mathfrak{X}_n$. With respect to a basis $\boldsymbol{x}_\kappa$, $\kappa = 1, \ldots, n$ the coordinate vector $\boldsymbol{x}$ of P is represented by

$$\boldsymbol{x} = \boldsymbol{x}_\kappa x^\kappa. \tag{5.1–7}$$

The numbers x^κ are called the *coordinates* of the point P with respect to the given origin and the given basis. Such a combination will be referred to as a coordinate system O, (κ), where (κ) stands for the vectors $\boldsymbol{x}_\kappa$, $\kappa = 1, \ldots, n$.

We wish to derive the equations of transformation when the coordinate system considered above is replaced by another O^*, (κ'). With reference to the equations (1.1–11) we readily find from (5.1–6)

$$x^\kappa = a^\kappa + x^{*\kappa'} q^\kappa_{\kappa'}, \tag{5.1–8}$$

the desired equations.

In differential geometry we are only concerned with transformations in exceptional cases, whereby also the origin is replaced. The most common case is, therefore, that with $\boldsymbol{a} = \boldsymbol{o}$. Then $\boldsymbol{x} = \boldsymbol{x}^*$ and the equations (5.1–8) reduce to

$$x^\kappa = x^{\kappa'} q^\kappa_{\kappa'}, \tag{5.1–9}$$

that is to say, the coordinates transform like the components of a vector.

5.2 – The metric point space

5.2.1 – Theorems on metric point spaces

A point space $\mathscr{R}$ associated with a metric vector space $\mathfrak{R}$ is called a *metric point space*. It is the natural generalization of ordinary euclidean space.

In a space of this kind we can define the angle between two lines, as the angle between the associated rays. It is usual to suppose that the generating vector of a ray is a unit vector. It is then called the *direction* of any line associated with the ray.

Lines are said to be *orthogonal* or *perpendicular* when their directions are orthogonal. More generally, a line is said to be orthogonal or perpendicular to a flat when its direction is orthogonal to all vectors of the vector space associated with the flat. The notion of orthogonality is independent of the orientation of the line. In some cases the line is called a *normal* of the flat.

It is also clear what must be understood by the principal angles between two finite dimensional flats in a metric point space.

Next we consider the important notion of *distance* between two points A and B in a metric point space, this being defined as the length of the vector $B-A$:

$$\text{dist}\, AB = |B-A|. \tag{5.2–1}$$

The distance has the following properties:

1) dist AB = dist BA,
2) dist $AB = 0$, if and only if A and B coincide,
3) dist $AB \leqq$ dist AC+dist BC.

Only the last statement, the *triangle inequality*, requires a proof. It is based on Schwarz's inequality (2.1–7).

By putting

$$z = B-A, \qquad x = B-C, \qquad y = C-A$$

we have

$$z = x+y$$

and

$$|z|^2 = |x+y|^2 = (x+y)(x+y) = |x|^2+2xy+|y|^2 \\ \leqq |x|^2+2|x||y|+|y|^2 = (|x|+|y|)^2,$$

whence

$$|z| \leqq |x|+|y|.$$

This completes the proof.

The triangle inequality reduces to an equality if and only if the points A, B and C are collinear, i.e., when they are on the same line. In fact, in that case the vectors $B-C$ and $C-A$ are linearly dependent.

5.2.2 – The metric space as a topological space

A set of points is called a *topological space* if a certain class of subspaces is assigned as a class of *open sets*, such that

1) the set itself and the empty set are open;
2) the intersection of two open sets is open;
3) the union of an arbitrary system of open subsets is an open subset.

Such a class of open sets is called a *topology*.

In a metric point space a set is said to be an open set when to any given point P of the set corresponds a positive number ε such that all points Q with dist $PQ < \varepsilon$ are included in the set. It is easily verified that the conditions above have been fulfilled. Hence a metric space is also a topological space. Thus it is possible to introduce the notion of continuity, the notion of a closed set, etc. For instance, a vector function $x(Q)$, which associates a vector in a metric vector space with every point Q of a subset of a metric space, is said to be *continuous* at a given point P when the following condition has been fulfilled: If ε denotes an arbitrary positive number then we can find a number δ such that $|x(Q)-x(P)| < \varepsilon$ for all points Q in the set with dist $PQ < \delta$.

A set is said to be *closed* when its complement, that is the set of all points not included in the set, is open.

A *continuous curve* in a metric space is the continuous map of a closed interval $a \leqq p \leqq b$. This means that to every number p of the interval corresponds a point $P(p)$ such that, given $\varepsilon > 0$, we can find a number δ with the property that dist $P(p)P(q) < \varepsilon$, provided $|p-q| < \delta$, where p is an arbitrary number of the interval and q included in the interval.

By a *region* we understand an open and connected subset of a metric point space. A set will be called *connected* if any two points A and B of the set can be connected by a continuous curve, the points of which are all included in the set.

5.2.3 – The number space

A particular case of the abstractly defined point space is the n-dimensional *number space*. We define a point of this space as a set of real numbers

$$\{q^1, \ldots, q^n\} \tag{5.2–2}$$

taken in order. That means that two points $\{p^1, \ldots, p^n\}$ and $\{q^1, \ldots, q^n\}$ coincide if and only if $p^\kappa = q^\kappa$, $\kappa = 1, \ldots, n$. The associated vector space $\mathfrak{X}_n$ consists of all n-tupels $(x^1, \ldots, x^n)$ considered in section 1.1.4. The relation between points and vectors is such that tot he points $\{p^1, \ldots, p^n\}$ and $\{q^1, \ldots, q^n\}$ corresponds the vector $(q^1-p^1, \ldots, q^n-p^n)$. It is at once clear that the conditions listed in section 5.1.1 have been fulfilled. In a number space it is customary to introduce a topology in the following way: A set is said to be open if to a point $P = \{p^1, \ldots, p^n\}$ of the set corresponds a number $\varepsilon > 0$ such that all points $Q = \{q^1, \ldots, q^n\}$ with $|p^\kappa - q^\kappa| < \varepsilon$, $\kappa = 1, \ldots, n$, belong to the set.

Next we consider two regions (κ) and (κ') in an n-dimensional number space. The points of the first region are denoted by $\{q^1, \ldots, q^n\}$, those of the second region by $\{q^{1'}, \ldots, q^{n'}\}$. The reason for the use of the same symbol (κ) for a region and a coordinate system already used earlier will be cleared up in section 5.3.1.

Next we suppose that we are given a set of n functions

$$q^{\kappa'} = q^{\kappa'}(q^1, \ldots, q^n), \qquad \kappa' = 1', \ldots, n' \tag{5.2–3}$$

defined throughout (κ) such that

1) they define a one-to-one correspondence between (κ) and (κ');

2) they possess all derivatives up to order u, $u > 0$, which are continuous throughout (κ);

3) at each point of (κ) the jacobian

$$\Delta = \frac{\partial(q^{1'}, \ldots, q^{n'})}{\partial(q^1, \ldots, q^n)} = \det \left[\frac{\partial q^{\kappa'}}{\partial q^{\kappa}}\right] \tag{5.2–4}$$

is different from zero. By well-known arguments it can be proved that the map of (κ) on (κ') defined by (5.2–3) is invertable, that is to say, we can find n functions

$$q^{\kappa} = q^{\kappa}(q^{1'}, \ldots, q^{n'}), \qquad \kappa = 1, \ldots, n, \tag{5.2–5}$$

defined throughout (κ) and satisfying the same conditions as listed above. When this situation occurs we shall say that the regions (κ) and (κ') are *equivalent* under a transformation of *class* C^u.

If (κ) and (κ') are equivalent as well as (κ') and (κ''), then (κ) and (κ'') are also equivalent, for let the equivalence between (κ) and (κ') be established by the functions $q^{\kappa'}(q^1, \ldots, q^n)$ and that between (κ') and (κ'') by the functions $q^{\kappa''}(q^{1'}, \ldots, q^{n'})$. When writing $q^{\kappa''}(q^1, \ldots, q^n)$ instead of $q^{\kappa''}(q^{1'}(q^1, \ldots, q^n), \ldots, q^{n'}(q^1, \ldots, q^n))$ we evidently have

$$\frac{\partial}{\partial q^{\kappa}} q^{\kappa''}(q^1, \ldots, q^n) = \frac{\partial}{\partial q^{\kappa'}} q^{\kappa''}(q^{1'}, \ldots, q^{n'}) \frac{\partial}{\partial q^{\kappa}} q^{\kappa'}(q^1, \ldots, q^n) \tag{5.2–6}$$

and the jacobian of the product correspondence is

$$\frac{\partial(q^{1''}, \ldots, q^{n''})}{\partial(q^1, \ldots, q^n)} = \frac{\partial(q^{1''}, \ldots, q^{n''})}{\partial(q^{1'}, \ldots, q^{n'})} \cdot \frac{\partial(q^{1'}, \ldots, q^{n'})}{\partial(q^1, \ldots, q^n)}. \tag{5.2–7}$$

5.2.4 – Functions on a number space

By a (real-valued) function on a number space we understand the mapping of a region of the number space into the set of real numbers. Let

$$\overset{(\kappa)}{\varphi}(q^1, \ldots, q^n) \tag{5.2–8}$$

denote a function defined throughout a region (κ). The function is said to be of *class* C^u when all derivatives up to the order u exists and are continuous throughout (κ). The first partial derivatives will be denoted by

$$\partial_{\lambda}\overset{(\kappa)}{\varphi} = \frac{\partial \overset{(\kappa)}{\varphi}}{\partial q^{\lambda}}, \tag{5.2–9}$$

the second derivatives by

$$\partial_{\lambda\mu}\overset{(\kappa)}{\varphi} = \frac{\partial^2 \varphi^{(\kappa)}}{\partial q^{\lambda}\,\partial q^{\mu}}, \tag{5.2–10}$$

etc.

If there is only one variable q we write

$$\mathrm{d}_q\, \varphi(q) \tag{5.2–11}$$

for the first derivative and

$$\mathrm{d}_q{}^2\varphi(q) \tag{5.2–12}$$

for the second derivative, etc.

Assume that the variables in (5.2–8) are replaced by n differentiable functions $q^1(q), \ldots, q^n(q)$ of a variable q. We shall write $\overset{(\kappa)}{\varphi}(q)$ instead of $\overset{(\kappa)}{\varphi}(q^1(q), \ldots, q^n(q))$. Then we have, on applying the chain rule

$$\mathrm{d}_q\overset{(\kappa)}{\varphi} = \partial_\kappa\overset{(\kappa)}{\varphi}\mathrm{d}_q q^\kappa. \tag{5.2–13}$$

Next we consider a function similar to $\overset{(\kappa)}{\varphi}$, also of class C^u defined throughout a region (κ'):

$$\overset{(\kappa')}{\varphi}(q^{1'}, \ldots, q^{n'}).$$

Suppose further that (κ) and (κ') are equivalent under a transformation (5.2–3) and that at corresponding points

$$\overset{(\kappa)}{\varphi}(q^1, \ldots, q^n) = \overset{(\kappa')}{\varphi}(q^{1'}, \ldots, q^{n'})$$

or, what amounts to the same,

$$\overset{(\kappa)}{\varphi}(q^1, \ldots, q^n) = \overset{(\kappa')}{\varphi}(q^{1'}(q^1, \ldots, q^n), \ldots, q^{n'}(q^1, \ldots, q^n))$$

throughout the region (κ). Then we shall say that the functions $\overset{(\kappa)}{\varphi}$ and $\overset{(\kappa')}{\varphi}$ are equal with respect to the correspondence (5.2–3) between (κ) and (κ').

In addition to these functions we also need *vectorial functions*, being the mapping of a region of the number space into a metric vector space. The notions of continuity and differentiability may be introduced in much the same way as in the case of ordinary functions. A vectorial function depending on the variables $q^1, \ldots, q^n$ will be denoted by

$$\overset{(\kappa)}{\boldsymbol{x}}(q^1, \ldots, q^n)$$

and its successive derivatives by $\partial_\lambda\overset{(\kappa)}{\boldsymbol{x}}$, $\partial_{\lambda\mu}\overset{(\kappa)}{\boldsymbol{x}}$, etc. The equality with respect to the correspondence (5.2–3) of two functions $\overset{(\kappa)}{\boldsymbol{x}}(q^1, \ldots, q^n)$ and $\overset{(\kappa')}{\boldsymbol{x}}(q^{1'}, \ldots, q^{n'})$ of class C^u is defined in just the same way as for ordinary functions.

By the symbol (κ) placed above the symbol denoting a function we indicated the region where the function is defined. In most cases confusion is not to be feared when we omit it. As a rule we shall write $\boldsymbol{x}(q^1, \ldots, q^n)$ instead of $\overset{(\kappa)}{\boldsymbol{x}}(q^1, \ldots, q^n)$, etc.

It should be noticed that the ordinary rules of calculus are valid in the case of vectorial functions. We recall the product rules

$$\partial_\lambda(\boldsymbol{x}\varphi) = \partial_\lambda \boldsymbol{x}\ \varphi + \boldsymbol{x}\partial_\lambda \varphi$$

and

$$\partial_\lambda(\boldsymbol{x}\boldsymbol{y}) = \boldsymbol{y}\partial_\lambda \boldsymbol{x} + \boldsymbol{x}\partial_\lambda \boldsymbol{y}.$$

5.3 - Manifolds

5.3.1 - DEFINITION

In a metric space $\mathscr{R}$ associated with a metric vector space $\mathfrak{R}$ we take a certain point O as origin. Then to every vector $\boldsymbol{x}$ of $\mathfrak{R}$ corresponds just one point of $\mathscr{R}$ and conversely. Hence we may describe point sets in $\mathscr{R}$ by means of sets of vectors in $\mathfrak{R}$.

Consider now a vector function $\boldsymbol{x}(q^1, \ldots, q^n)$ defined throughout a region (κ) in an n-dimensional number space. It defines a mapping of this region on a point set in the metric point space, called an *n-dimensional manifold of class* κ, when κ is the class of the given vector function. It may happen that another mapping defined throughout a region (κ') yields the same point set. This will be considered as the same manifold as considered above, when the defining functions $\boldsymbol{x}(q^1, \ldots, q^n)$ and $\boldsymbol{x}(q^{1'}, \ldots, q^{n'})$ are of the same class and equal with respect to the correspondence (5.2–5) between (κ) and (κ'). By a *point of a manifold* we understand a point of the region (κ) together with its image in $\mathscr{R}$. Since the image is determined unambiguously by the coordinates $q^1, \ldots, q^n$ we shall briefly speak of the point q^κ, $\kappa = 1, \ldots, n$, with the understanding that q^κ in (κ) and $q^{\kappa'}$ in (κ') denote the same point of the manifold, when they are equivalent under the correspondence (5.2–5) defining the equality of manifolds. The numbers q^κ are often called *parameters* of the point.

Only those properties of manifolds are of interest which are un-influenced by a change of parameters. More correctly stated: by an *allowable change*, that is to say, by replacing q^κ by $q^{\kappa'}$ corresponding with q^κ in the correspondence (5.2–5) defining the equality of manifolds.

Thus, for instance, the notions *dimension* and *class* have a meaning, for they are the same for a class of mutually equal manifolds.

Another important concept is the following. Let Q denote a point of the manifold with parameters q^κ. By a *neighbourhood* of Q on the manifold we understand a set of points P whose parameters p^κ indicate points of (κ) which vary throughout an open set of (κ) including the point with coordinates q^κ. It is easily shown that this concept does not depend on a particular choice of the parameter region (κ).

Differential geometry is only concerned with *local properties* of a manifold; those are properties which may be stated by means of a sufficiently small neighbourhood of a point. Thus, for instance, differentiability of the defining vector functions $\boldsymbol{x}(q^\kappa)$ is such a property.

Up to now the notion of manifold is too general to yield interesting properties. We therefore make some restrictions by confining ourselves to the so-called regular points. A point of a manifold is called a *regular point,* when at this point the vectors

$$\partial_1\boldsymbol{x}, \ldots, \partial_n\boldsymbol{x} \tag{5.3–1}$$

are linearly independent. On account of (5.2–4) and the fact that

$$\boxed{\partial_{\kappa'}\boldsymbol{x} = \partial_\kappa\boldsymbol{x}\,\partial_{\kappa'}q^\kappa,} \qquad \kappa' = 1', \ldots, n' \tag{5.3–2}$$

the vectors $\partial_{\kappa'}\boldsymbol{x}$ are also linearly independent. Hence the condition of regularity does not depend on a particular choice of the parameters.

5.3.2 – The tangent space

The vectors (5.3–1) evaluated at a given point of the manifold span an n-dimensional vector space, called the *tangent space* at the given point. On account of (5.3–2) the vectors $\partial_{\kappa'}\boldsymbol{x}$ span the same space. The vectors (5.3–1) may be taken as the elements of a basis in the tangent space. In many cases it is convenient to write $\boldsymbol{x}_\kappa$ instead of $\partial_\kappa\boldsymbol{x}$ and to introduce the reciprocal basis $\boldsymbol{x}^\kappa$. Hence, to every parameter region (κ) corresponds a frame in the tangent space and it is natural to denote this frame also by (κ). This is in accordance with the agreement made in section 2.2.2. Writing also $q^\kappa_{\kappa'}$ instead of $\partial_{\kappa'}q^\kappa$ we see that (5.3–2) takes the form (1.1–11) and thus we may state: *A change of parameters induces a change of basis in every tangent space.*

A particular case occurs when the dimension of the manifold is one. Such a manifold is called a *curve* and is defined by a vector function $\boldsymbol{x}(q)$, where q varies throughout an open interval $a < q < b$, with the additional condition of regularity, expressing that the derivative (being continuous throughout the interval) is nowhere equal to the zero vector. If $p(q)$ is a continuously differentiable function of q and $\mathrm{d}_q p \neq 0$ throughout the interval then p describes also an open interval and may be taken as a parameter. Since

$$\mathrm{d}_q\boldsymbol{x} = \mathrm{d}_p\boldsymbol{x}\,\mathrm{d}_q p \tag{5.3–3}$$

the vector $\mathrm{d}_p\boldsymbol{x} \neq \boldsymbol{o}$ throughout the range of q. The ray spanned by a tangent vector $\mathrm{d}_q\boldsymbol{x}$ is called the *tangent* of the curve at the given point.

Next we replace the parameter q^κ, $\kappa = 1, \ldots, n$, of the points of an n-dimensional manifold by functions $q^\kappa(q)$ which have continuous derivatives throughout a certain interval which do not vanish simultaneously. By writing

$$x(q) = x(q^1(q), \ldots, q^n(q)) \tag{5.3-4}$$

we have obtained a function of the parameter q which defines a curve, whose points are included in the manifold. The tangent vector at a given point is evidently

$$d_q x = \partial_\kappa x \, d_q q^\kappa \neq o \tag{5.3-5}$$

and belongs to the tangent space at the point under consideration. Hence:

The tangents at a given point of all curves on the manifold passing through this point are included in the tangent space of the manifold at this point.

5.3.3 – The metric tensor

Let us introduce the functions

$$g_{\lambda\mu} = \partial_\lambda x \, \partial_\mu x, \qquad \lambda, \mu = 1, \ldots, n. \tag{5.3-6}$$

A change of parameters gives rise to functions $g_{\lambda'\mu'}$ such that

$$g_{\lambda'\mu'} = g_{\lambda\mu} \partial_{\lambda'} q^\lambda \partial_{\mu'} q^\mu. \tag{5.3-7}$$

Hence in every tangent space they transform like the components of the metric tensor. Actually the $g_{\lambda\mu}$ are functions on the manifold and their values at a given point are the components of a tensor. Such a system of tensors is often called a *tensor field*, but we shall refer to this as a tensor.

Assuming that the class of the manifold is at least two, the functions are of the class $\geqq 1$. On solving the equations

$$x^\lambda g_{\lambda\mu} = \partial_\mu x = x_\mu \tag{5.3-8}$$

we see, that then the x^λ are also functions of the same class and the same is true for the functions

$$g^{\lambda\mu} = x^\lambda x^\mu. \tag{5.3-9}$$

Thus we see that every tangent space is a metric vector space. We shall express this by saying that a manifold possesses at each of its points a *local metric*.

The tensor considered in this section is called the *metric tensor on the manifold*.

5.3.4 – Correspondences between two manifolds

In the same metric point space we consider two manifolds defined by the functions $x(q^\kappa)$ and $y(p^\kappa)$, $\kappa = 1, \ldots, n$, respectively. As is usual $x(q^\kappa)$ stands for $x(q^1, \ldots, q^n)$. The first manifold is the set of points Q, such that

$$Q - O = x(q^\kappa). \tag{5.3-10}$$

The second manifold is the set of points P, such that

$$P - O = y(p^\kappa). \tag{5.3-11}$$

We assume that the points $\{q^1, \ldots, q^n\}$ in the number space range through a region (κ) and the points $\{p^1, \ldots, p^n\}$ through a region (λ).

Next we assume that we are given a set of n functions $p^\kappa(q^1, \ldots, q^n)$ defined throughout (κ). Suppose that both manifolds are of class C^u. These functions are also assumed to be of class C^u. When these functions are such that they establish a one-to-one correspondence between (κ) and (λ) and when, moreover, throughout (κ) the condition

$$\frac{\partial(p^1, \ldots, p^n)}{\partial(q^1, \ldots, q^n)} \neq 0 \tag{5.3-12}$$

has been fulfilled, we shall say that the two manifolds are in a *regular correspondence*. This correspondence is locally one-to-one and bicontinuous.

On account of the conditions stated we may take the variables q^κ as parameters for the second manifold. Then this second manifold is defined by a vector function $\overset{*}{\boldsymbol{x}}(q^1, \ldots, q^n)$ which stands for $\boldsymbol{y}(p^1(q^1, \ldots, q^n), \ldots, p^n(q^1, \ldots, q^n))$. From now on we shall always assume that two manifolds in a regular correspondence are referred to the same system of parameters.

The metric tensor on the second manifold is given by

$$\overset{*}{g}_{\lambda\mu} = \partial_\lambda \overset{*}{\boldsymbol{x}} \, \partial_\mu \overset{*}{\boldsymbol{x}} \tag{5.3-13}$$

and since these functions $\overset{*}{g}_{\lambda\mu}$ are defined throughout (κ) we may consider them as the components of a (symmetric) tensor on the first manifold, of course not the metric tensor.

By corresponding vectors at corresponding points we understand vectors $\boldsymbol{a}$ and $\overset{*}{\boldsymbol{a}}$ such that

$$\boldsymbol{a} = \boldsymbol{x}_\lambda a^\lambda, \qquad \overset{*}{\boldsymbol{a}} = \overset{*}{\boldsymbol{x}}_\lambda a^\lambda, \tag{5.3-14}$$

where $\boldsymbol{x}_\lambda = \partial_\lambda \boldsymbol{x}$, $\overset{*}{\boldsymbol{x}}_\lambda = \partial_\lambda \overset{*}{\boldsymbol{x}}$, evaluated for the same values of the parameters. In particular we take for $\boldsymbol{a}$ a unit vector $\boldsymbol{u}$ on the first manifold. It corresponds with a vector $\overset{*}{\boldsymbol{u}}$ on the second manifold of norm

$$\omega = \overset{*}{g}_{\lambda\mu} u^\lambda u^\mu. \tag{5.3-15}$$

This number is called the *distortion* on the first manifold in the direction $\boldsymbol{u}$.

We may picture the correspondence by means of the linear operator, the distortion operator,

$$w_\lambda = \overset{*}{g}_{\lambda\mu} u^\mu, \tag{5.3-16}$$

which associates to a vector $\boldsymbol{u}$ a vector $\boldsymbol{w}$ on the first manifold. In general the directions of $\boldsymbol{u}$ and $\boldsymbol{w}$ are different, but it may happen that they are the same, viz., when $\boldsymbol{u}$ is an eigenvector of the operator with components $\overset{*}{g}_{\lambda\mu}$. They indicate *principal directions for the distortion* and the corresponding eigenvalues are called the *principal distortions*. They are the roots of the equation

$$\det\,[\overset{*}{g}_{\lambda\mu}-\omega g_{\lambda\mu}] = 0. \tag{5.3–17}$$

Referring to the first theorem of section 3.2.3 we may state *Tissot's theorem*:

At every point of a manifold which is in regular correspondence with a second manifold there is a system of n mutually orthogonal principal directions of distortion, being the principal directions of the distortion operator.

When all principal distortions are equal, then

$$\overset{*}{g}_{\lambda\mu} = \lambda g_{\lambda\mu}. \tag{5.3–18}$$

In this case the angle between two vectors $\boldsymbol{u}$, $\boldsymbol{v}$ and their corresponding vectors $\overset{*}{\boldsymbol{u}}$, $\overset{*}{\boldsymbol{v}}$ are equal. A point where this property holds is called an *isotropic point* for the correspondence. When all points of the manifold are isotropic then the correspondence is *conformal*. In the particular case that $\lambda = 1$ we have an *isometric correspondence* since then the length of a vector is equal to the length of the corresponding vector.

5.4 – Scalar invariants

5.4.1 – The gradient of a scalar

Suppose with every parameter system (κ) of an n-dimensional manifold is associated a function $\overset{(\kappa)}{\varphi}$ such that the functions for corresponding parameter systems in allowable transformations are equal. Then the $\overset{(\kappa)}{\varphi}$ may be considered as representatives of a function φ defined throughout the manifold independent of a particular choice of the parameter system. We may also express this by saying that the $\overset{(\kappa)}{\varphi}$ are the components in the various parameter systems of a tensor field of zero valency, called a *scalar invariant*.

Let $q^\kappa(q)$, $\kappa = 1, \ldots, n$, define a curve on the manifold. Assuming that φ is of class C^u, $u \geqq 1$, the function

$$\varphi(q) = \overset{(\kappa)}{\varphi}(q^1(q), \ldots, q^n(q)) \tag{5.4–1}$$

is continuously differentiable with respect to q. By the *rate of change* of φ along the curve is understood the derivative

$$\mathrm{d}_q\varphi = \partial_\kappa\varphi\,\mathrm{d}_q q^\kappa. \tag{5.4–2}$$

This value depends not only on the direction of the curve at the given point but also on the parameter q. It is, therefore, natural to replace $\mathrm{d}_q q^\kappa$ by the components u^κ of a unit vector $\boldsymbol{u}$ denoting the direction of the curve and to consider the number

$$f(\boldsymbol{u}) = u^\kappa\partial_\kappa\varphi, \tag{5.4–3}$$

called the *derivative of* φ *in the direction* $\boldsymbol{u}$.

The partial derivatives $\partial_\kappa\varphi$ *are the covariant components of a vector,* the *gradient* of φ. For a change of parameters induces the transformation

$$\partial_{\kappa'}\varphi = \partial_\kappa\varphi\,\partial_{\kappa'}q^\kappa. \tag{5.4–4}$$

This vector will be denoted by grad φ or $\nabla\varphi$ (nabla φ), i.e.,

$$\boxed{\operatorname{grad}\varphi = \nabla\varphi = \boldsymbol{x}^\kappa\,\partial_\kappa\varphi.} \tag{5.4–5}$$

The derivative of φ in the direction $\boldsymbol{u}$, being the expression (5.4–3), may be written as

$$f(\boldsymbol{u}) = \boldsymbol{u}\nabla\varphi. \tag{5.4–6}$$

If θ denotes the angle between the vectors $\boldsymbol{u}$ and $\nabla\varphi$ then

$$f(\boldsymbol{u}) = |\nabla\varphi|\cos\theta, \tag{5.4–7}$$

assuming that $\nabla\varphi$ is not the zero vector. Under this assumption we have:

The derivative of the scalar invariant φ *at a given point takes its maximal value when the direction* $\boldsymbol{u}$ *coincides with the direction of* $\nabla\varphi$, *that is when*

$$\boldsymbol{u}|\nabla\varphi| = \nabla\varphi \tag{5.4–8}$$

The square of the expression (5.4–3) is the quadratic form

$$\partial_\lambda\varphi\,\partial_\mu\varphi\;u^\lambda u^\mu. \tag{5.4–9}$$

It is easy to prove that the rank of this form is unity, provided $\nabla\varphi \neq 0$ and that an eigenvector belonging to the non-zero eigenvalue is the vector $\nabla\varphi$.

By *Beltrami's differential parameter of the first order* of two scalars φ and ψ is understood the scalar product of their gradients. It is usually denoted by $\Delta_1(\varphi, \psi)$, i.e.,

$$\boxed{\Delta_1(\varphi, \psi) = \nabla\varphi\nabla\psi = g^{\lambda\mu}\,\partial_\lambda\varphi\,\partial_\mu\psi.} \tag{5.4–10}$$

In particular, $\Delta_1(\varphi, \varphi)$ is the norm of grad φ.

Consider the scalar, which is equal to q^κ in the system (κ), where the superscript κ denotes a fixed natural number. In any other system (κ') this scalar is represented by the function $q^\kappa(q^{1'}, \ldots, q^{n'})$ occurring in the equations of transformation from (κ') to (κ). Since $\partial_\lambda q^\kappa = \delta^\kappa_\lambda$, we readily find from (5.4–10):

$$g^{\lambda\mu} = \triangle_1(q^\lambda, q^\mu), \tag{5.4–11}$$

being a scalar which is equal to $g^{\lambda\mu}$ in the system (κ) and to $g^{\lambda'\mu'}\partial_{\lambda'} q^\lambda \partial_{\mu'} q^\mu$ in the system (κ').

5.4.2 – The divergence of a vector

Suppose that to every point in a neighbourhood of a given point on a manifold corresponds a tangential vector $\boldsymbol{a}$. The set of all these vectors constitutes a *vector field*. We further suppose that this field is of class one at least. That means that the components with respect to a parameter system (κ) are continuously differentiable functions of the variables q^κ and it is easily checked, on account of the equations

$$a^{\lambda'} = a^\lambda \partial_\lambda q^{\lambda'}, \tag{5.4–12}$$

that this property holds for every parametersystem. On manifolds we are almost always concerned with fields. But instead of vector field we prefer to use the brief name vector.

It is our aim to describe a process which yields a scalar invariant from a vector given by its contravariant components. This requires some preliminary considerations.

The covariant components of the metric tensor transform according to the equations

$$g_{\lambda'\mu'} = g_{\lambda\mu}\partial_{\lambda'} q^\lambda \partial_{\mu'} q^\mu. \tag{5.4–13}$$

The gramian of the vectors $\partial_\kappa \boldsymbol{x}$, $\kappa = 1, \ldots, n$, is $\det[g_{\lambda\mu}]$. On applying the product rule for determinants we find

$$\det[g_{\lambda'\mu'}] = \det[g_{\lambda\mu}] \det{}^2[\partial_{\kappa'} q^\kappa].$$

Next we observe that

$$\det[\partial_{\kappa'} q^\kappa] = \Delta^{-1},$$

where Δ is the Jacobian (5.2–6). Writing

$$\overset{(\kappa)}{g} = \det[g_{\lambda\mu}]$$

we evidently have

$$\boxed{\overset{(\kappa')}{g} = \overset{(\kappa)}{g}\Delta^{-2}.} \tag{5.4–14}$$

Thus it appears that the functions $\overset{(\kappa)}{g}$ do not define a scalar invariant. More generally we may consider a set of functions $\overset{(\kappa)}{a}$ corresponding to the parameter systems (κ) such that

$$\overset{(\kappa')}{a} = \overset{(\kappa)}{a} \Delta^{-w}. \tag{5.4–15}$$

Then we shall say that $\overset{(\kappa)}{a}$ is the component of a *density of weight w* with respect to the system (κ). Hence $\overset{(\kappa)}{g}$ represents a density of weight two.

We wish to investigate more closely a density of weight equal to one. We are particularly interested in the first derivatives in order to obtain an entity analogous to a gradient. By differentiating the relation

$$\overset{(\kappa')}{a} = \overset{(\kappa)}{a} \Delta^{-1} \tag{5.4–16}$$

we find

$$\partial_{\lambda'} \overset{(\kappa')}{a} = \partial_\lambda \overset{(\kappa)}{a} \, \partial_{\lambda'} q^\lambda \Delta^{-1} - \overset{(\kappa)}{a} \, \partial_\lambda \Delta \, \partial_{\lambda'} q^\lambda \Delta^{-2}. \tag{5.4–17}$$

This is a rather unpleasant law of transformation. It is, however, possible to express $\partial_\lambda \Delta$ in terms of derivatives of the functions $q^{\kappa'}$. To this end we first observe that

$$\frac{\partial \Delta}{\partial q^{\kappa'}_\mu} \tag{5.4–18}$$

is the algebraic complement of the element $q^{\kappa'}_\kappa$ in the determinant Δ, where $q^{\kappa'}_\kappa$ stands for $\partial_\kappa q^{\kappa'}$. Hence, by virtue of a well-known theorem on determinants,

$$q^{\kappa'}_\lambda \frac{\partial \Delta}{\partial q^{\kappa'}_\mu} = \delta^\mu_\lambda \Delta,$$

whence

$$\frac{\partial \Delta}{\partial q^{\kappa'}_\mu} = \delta^\mu_\lambda q^\lambda_{\kappa'} \Delta = q^\mu_{\kappa'} \Delta. \tag{5.4–19}$$

On applying the chain-rule we readily find

$$\partial_\lambda \Delta = \frac{\partial \Delta}{\partial q^{\kappa'}_\mu} \partial_\lambda q^{\kappa'}_\mu = \partial_{\lambda\mu} q^{\kappa'} \partial_{\kappa'} q^\mu \Delta. \tag{5.4–20}$$

At first sight we have gained nothing. But a similar expression as that on the right of (5.4–20) occurs when we write down the law of transformation of the functions $\partial_\lambda v^\lambda$, where v^λ are the components of a covariant vector. In fact, differentiating both members of the equation

$$v^{\mu'} = v^{\mu}\,\partial_{\mu} q^{\mu'} \tag{5.4–21}$$

we get

$$\partial_{\lambda'} v^{\mu'} = \partial_{\lambda} v^{\mu}\,\partial_{\lambda} q^{\lambda'}\,\partial_{\mu} q^{\mu'} + v^{\mu}\,\partial_{\lambda\mu} q^{\mu'}\,\partial_{\lambda'} q^{\lambda}.$$

On equating the indices λ' and μ' and performing the summation according to the summation convention, we find

$$\begin{aligned}\partial_{\lambda'} v^{\lambda'} &= \partial_{\lambda} v^{\lambda} + v^{\mu}\,\partial_{\lambda\mu} q^{\lambda'}\,\partial_{\lambda'} q^{\lambda}\\ &= \partial_{\lambda} v^{\lambda} + v^{\lambda}\,\partial_{\lambda\mu} q^{\kappa'}\,\partial_{\kappa'} q^{\mu},\end{aligned}$$

and this may also be written as

$$\partial_{\lambda'} v^{\lambda'} = \partial_{\lambda} v^{\lambda} + v^{\lambda}\,\partial_{\lambda}\Delta\,\Delta^{-1}, \tag{5.4–22}$$

where we have made use of (5.4–20). Comparing this with (5.4–17) it is natural to eliminate $\partial_{\lambda}\Delta$. When multiplying both members of (5.4–17) by $v^{\lambda'}$ we get

$$v^{\lambda'}\,\partial_{\lambda'} \overset{(\kappa')}{a} = v^{\lambda}\,\partial_{\lambda} \overset{(\kappa)}{a} - \overset{(\kappa)}{a} v^{\lambda}\,\partial_{\lambda}\Delta\,\Delta^{-2};$$

multiplying both members (5.4–22) by $\overset{(\kappa')}{a}$ we obtain

$$\overset{(\kappa')}{a}\partial_{\lambda'} v^{\lambda'} = \overset{(\kappa)}{a}\partial_{\lambda} v^{\lambda}\Delta^{-1} + \overset{(\kappa)}{a} v^{\lambda}\,\partial_{\lambda}\Delta\,\Delta^{-2}.$$

On adding corresponding members of the results obtained we arrive at

$$\partial_{\lambda'}(\overset{(\kappa')}{a} v^{\lambda'}) = \partial_{\lambda}(\overset{(\kappa)}{a} v^{\lambda})\Delta^{-1}, \tag{5.4–23}$$

that is to say, at a density of weight unity. Now we always have a density of weight unity at our disposal, viz. the density $\overset{(\kappa)}{\sqrt{g}} = \sqrt{g}$. Replacing $\overset{(\kappa)}{a}$ by this density and dividing by $\sqrt{g}$, we see that

$$\boxed{\operatorname{div} \mathbf{v} = \frac{1}{\sqrt{g}}\,\partial_{\lambda}(\sqrt{g}\,v^{\lambda})} \tag{5.4–24}$$

is a scalar invariant, called the *divergence* of the vector $\mathbf{v}$.

The n-dimensional point space may be considered as an n-dimensional manifold homeomorphic to the number space. In fact, introducing a basis $\mathbf{x}_{\kappa}$, $\kappa = 1, \ldots, n$ of constant vectors we may associate with the point q^{κ}, $\kappa = 1, \ldots, n$, in the number space the vector $\mathbf{x}_{\kappa} q^{\kappa}$. It is clear that in this case g is a constant and with respect to this system of parameters (5.4–24) reduces to

$$\operatorname{div} \mathbf{v} = \partial_{\lambda} v^{\lambda},$$

the well-known expression in elementary vector calculus.

If φ is a scalar invariant and $\boldsymbol{v}$ a vector, we may express the divergence of $\varphi\boldsymbol{v}$ in terms of the divergence of $\boldsymbol{v}$. A simple computation yields at once the result

$$\operatorname{div} \varphi\boldsymbol{v} = \varphi \operatorname{div} \boldsymbol{v} + \boldsymbol{v}\nabla\varphi. \tag{5.4–25}$$

5.4.3 – THE LAPLACIAN OF A SCALAR INVARIANT

We wish to evaluate the divergence of the gradient of a scalar φ. Since in (5.4–24) the contravariant components of the vector $\boldsymbol{v}$ occur, we must insert the contravariant components of the gradient, these being

$$g^{\lambda\mu}\,\partial_\mu\,\varphi.$$

Thus by (5.4–24)

$$\boxed{\operatorname{div}\ \operatorname{grad}\varphi = \frac{1}{\sqrt{g}}\,\partial_\lambda(\sqrt{g}\,g^{\lambda\mu}\partial_\mu\varphi).} \tag{5.4–26}$$

In the case of an n-dimensional space referred to an orthonormal basis the expression on the right reduces to

$$\partial_1{}^2\varphi + \ldots + \partial_n{}^2\varphi,$$

the laplacian of φ. For this reason (5.4–26) is also called the *laplacian of the scalar invariant* φ. It is usually denoted by $\triangle\varphi$ or $\nabla^2\varphi$. It was introduced in this general form by Beltrami and called by him *a differential parameter of the second order*.

If φ and ψ are scalar invariants, then

$$\nabla(\varphi\psi) = (\nabla\varphi)\psi + (\nabla\psi)\varphi. \tag{5.4–27}$$

Taking the divergence of both members and making use of (5.4–25) we obtain

$$\triangle(\varphi\psi) = \varphi\triangle\psi + 2\nabla\varphi\nabla\psi + \psi\triangle\varphi.$$

The equation $\triangle\varphi = 0$ is a partial differential equation of the second order. Any solution having continuous derivatives up to at least the second order is called a *harmonic function* on the manifold.

CHAPTER 6

CURVES

The simplest manifolds are the one dimensional manifolds, the curves. They are characterized by a vector function $\boldsymbol{x}(q)$ depending on one variable. We shall denote the dimension of the space by $r+1$, where $r > 0$ and we wish to assume that the class of $\boldsymbol{x}(q)$ is C^{r+1} at least. Most theorems about curves centre around a set of equations, Frenet's equations, which are fundamental in the theory of curves. The proof of some existence theorems is based on the well-known existence theorem about systems of differential equations of the first order. We assume this theorem to be known.

6.1 – Elementary notions

6.1.1 – OSCULATING SPACES

Let us consider a curve defined by the vector function

$$\boldsymbol{x}(q), \tag{6.1–1}$$

depending on the variable q. We are only concerned with the points where

$$\mathrm{d}_q \boldsymbol{x} \neq \boldsymbol{o}, \tag{6.1–2}$$

the regular points of the curve.

Assuming that for a certain value of q and a number h

$$\mathrm{d}_q \boldsymbol{x},\ \mathrm{d}_q{}^2 \boldsymbol{x}, \ldots, \mathrm{d}_q{}^{h+1} \boldsymbol{x} \tag{6.1–3}$$

are linearly independent, they span an h-dimensional vector space, called *the h-dimensional osculating space* of the curve at the point corresponding to the given value of the parameter.

It is easy to see that this space does not depend on the particular choice of the parameter along the curve. In fact, let $q(p)$ denote a function of p with $\mathrm{d}_p q \neq 0$ throughout the interval of p. Then

$$\begin{aligned} \mathrm{d}_p \boldsymbol{x} &= \mathrm{d}_q \boldsymbol{x}\, \mathrm{d}_p q, \\ \mathrm{d}_p{}^2 \boldsymbol{x} &= \mathrm{d}_q \boldsymbol{x}\, \mathrm{d}_p{}^2 q + \mathrm{d}_q{}^2 \boldsymbol{x} (\mathrm{d}_p q)^2 \\ &\cdots\cdots\cdots\cdots \\ \mathrm{d}_p{}^{h+1} \boldsymbol{x} &= \mathrm{d}_q \boldsymbol{x}\, \mathrm{d}_p{}^{h+1} q + \ldots + \mathrm{d}_q{}^{h+1} \boldsymbol{x} (\mathrm{d}_p q)^{h+1}. \end{aligned} \tag{6.1–4}$$

This is a linear transformation of the set of vectors $\mathrm{d}_q \boldsymbol{x}, \ldots, \mathrm{d}_q^{h+1} \boldsymbol{x}$, the determinant of the transformation matrix having the value

$$(\mathrm{d}_p q)^{1+2+\cdots+(h+1)} = (\mathrm{d}_p q)^{\frac{1}{2}(h+1)(h+2)} \neq 0.$$

This proves the assertion.

The one-dimensional osculating space is the tangent. By virtue of (6.1–2) the tangent always exists.

Let us now suppose that the vectors (6.1–3) are linearly dependent throughout the interval where q varies, h being a fixed number. Then we may state:

When the vectors (6.1-3) *are linearly dependent for* $a < q < b$, *then the curve is included in an h-dimensional linear subspace.*

Bij hypothesis there are $h+1$ functions $\underset{0}{\alpha}(q), \ldots, \underset{h}{\alpha}(q)$ such that

$$\mathrm{d}_q \boldsymbol{x}\underset{0}{\alpha}(q) + \ldots + \mathrm{d}_q^{h+1} \boldsymbol{x}\underset{h}{\alpha}(q) = \boldsymbol{o}$$

or, if we put $\boldsymbol{y} = \mathrm{d}_q \boldsymbol{x}$:

$$\boldsymbol{y}\underset{0}{\alpha}(q) + \ldots + \mathrm{d}_p^h \boldsymbol{y}\underset{h}{\alpha}(q) = \boldsymbol{o}. \tag{6.1–5}$$

On multiplying both members by a constant vector $\boldsymbol{r}$ we see that every solution of the differential equation (6.1–5) for $\boldsymbol{y}$ gives rise to a solution of the homogeneous linear differential equation

$$y\underset{0}{\alpha}(q) + \ldots + \mathrm{d}_q^h y\underset{h}{\alpha}(q) = 0. \tag{6.1–6}$$

This equation possesses the system of solutions

$$y = c_1 \eta^1 + \ldots + c_h \eta^h = c_\kappa \eta^\kappa$$

and, consequently, the general solution of (6.1–5) is

$$\boldsymbol{y} = \boldsymbol{c}_\kappa \eta^\kappa, \tag{6.1–7}$$

whence

$$\boldsymbol{x} = \boldsymbol{c}_\kappa \xi^\kappa$$

with

$$\xi^\kappa = \int \eta^\kappa \mathrm{d}q, \qquad \kappa = 1, \ldots, h.$$

Thus we see that $\boldsymbol{x}$ varies throughout a fixed h-dimensional vector space.

6.1.2 – The arc length

For many purposes it is convenient to introduce a parameter s with the property that

$$\mathrm{d}_s \boldsymbol{x} \tag{6.1–8}$$

is a unit vector. Under this assumption it follows from

$$\mathrm{d}_q \boldsymbol{x} = \mathrm{d}_s \boldsymbol{x}\, \mathrm{d}_q s$$

that

$$\boxed{(\mathrm{d}_q s)^2 = \mathrm{d}_q \boldsymbol{x}\, \mathrm{d}_q \boldsymbol{x}.} \tag{6.1–9}$$

Hence a function s satisfying the condition that (6.1–8) is a unit vector, is given by a solution s of the differential equation

$$\mathrm{d}_q s = \sqrt{\mathrm{d}_q \boldsymbol{x}\, \mathrm{d}_q \boldsymbol{x}}.$$

This solution is determined up to an additive constant. All desired parameters are representable as $\pm s+c$.

The parameter s may be interpreted geometrically as the arc length measured along the curve.

Differentiating the equation

$$\mathrm{d}_s \boldsymbol{x}\, \mathrm{d}_s \boldsymbol{x} = 1 \tag{6.1–10}$$

we get

$$\mathrm{d}_s \boldsymbol{x}\, \mathrm{d}_s^2 \boldsymbol{x} = 0. \tag{6.1–11}$$

Hence $\mathrm{d}_s^2 \boldsymbol{x}$ *is orthogonal to* $\mathrm{d}_s \boldsymbol{x}$.

The vector $\mathrm{d}_s^2 \boldsymbol{x}$ *is the zero vector if and only if* $\mathrm{d}_s \boldsymbol{x}$ *and* $\mathrm{d}_s^2 \boldsymbol{x}$ *are linearly dependent.*

The assertion is trivial when $\mathrm{d}_s^2 \boldsymbol{x} = \boldsymbol{o}$. Let us assume that there are functions $\underset{0}{\alpha}(s)$ and $\underset{1}{\alpha}(s)$ such that

$$\mathrm{d}_s \boldsymbol{x}\, \underset{0}{\alpha} + \mathrm{d}_s^2 \boldsymbol{x}\, \underset{1}{\alpha} = \boldsymbol{o}. \tag{6.1–12}$$

The function $\underset{1}{\alpha}$ cannot take the value zero, for $\mathrm{d}_s \boldsymbol{x} \neq \boldsymbol{o}$. On multiplying both members of (6.1–12) by $\mathrm{d}_s \boldsymbol{x}$ we see, by virtue of (6.1–10) and (6.1–11), that $\underset{0}{\alpha} = 0$ and this proves the assertion. *The equation*

$$\mathrm{d}_s^2 \boldsymbol{x} = 0 \tag{6.1–13}$$

characterizes a straight line.

When $\mathrm{d}_s^2 \boldsymbol{x} \neq 0$ then we may find a unit vector $\underset{1}{\boldsymbol{u}}$, such that

$$\boxed{\mathrm{d}_s^2 \boldsymbol{x} = \underset{1}{\boldsymbol{u}}\kappa,} \qquad \kappa > 0. \tag{6.1–14}$$

The uniquely determined vector $\underset{1}{\boldsymbol{u}}$ is called the *normal vector* of the curve and κ is the *curvature*. The vector $\mathrm{d}_s^2 \boldsymbol{x}$ is known as the *curvature vector*.

It may happen that in isolated points the relation (6.1–13) holds. Then in general $\underset{1}{\boldsymbol{u}}$ can be defined by continuity and (6.1–14) is also valid provided we take $\kappa = 0$. These points are called *flexes* of the curve.

6.2 – Frenet's equations

6.2.1 – Frenet's equations

We consider the vectors (assumed to be linearly independent)

$$\mathrm{d}_s\boldsymbol{x},\ \mathrm{d}_s^2\boldsymbol{x},\ \ldots,\ \mathrm{d}_s^{r+1}\boldsymbol{x}, \tag{6.2–1}$$

associated with the curve represented by the vector function $\boldsymbol{x}(s)$, referred to the arc length as parameter. On applying the Gram-Schmidt process described in section 2.3.1, starting with $\mathrm{d}_s\boldsymbol{x}$, we obtain an orthonormal set

$$\underset{0}{\boldsymbol{u}}(s),\ \ldots,\ \underset{r}{\boldsymbol{u}}(s) \tag{6.2–2}$$

such that $\underset{h}{\boldsymbol{u}}(s)$ is a linear combination of $\mathrm{d}_s\boldsymbol{x}, \ldots, \mathrm{d}_s^{h+1}\boldsymbol{x}$, while $\mathrm{d}_s^{h+1}\boldsymbol{x}$ is linearly dependent on $\underset{0}{\boldsymbol{u}}(s), \ldots, \underset{h}{\boldsymbol{u}}(s)$. Since the orthogonalization process is performed by rational operations we may infer that the vectors (6.2–2) have the same differentiability properties as the original vectors (6.2–1).

Next we write

$$\mathrm{d}_s\underset{h}{\boldsymbol{u}} = \sum_{k=0}^{r} \underset{k}{\boldsymbol{u}}\ \underset{kh}{\alpha} \qquad h = 0, \ldots, r \tag{6.2–3}$$

and observe that $\underset{kh}{\alpha} = 0$ whenever $k > h+1$. In fact, $\underset{h}{\boldsymbol{u}}$ is included in the space spanned by $\mathrm{d}_s\boldsymbol{x}, \ldots, \mathrm{d}_s^{h+1}\boldsymbol{x}$. Hence $\mathrm{d}_s\underset{h}{\boldsymbol{u}}$ is a vector of the space spanned by $\mathrm{d}_s\boldsymbol{x}, \ldots, \mathrm{d}_s^{h+2}\boldsymbol{x}$, that is the space spanned by $\underset{0}{\boldsymbol{u}}, \ldots, \underset{h+1}{\boldsymbol{u}}$. It follows that in (6.2–3) the vectors $\underset{k}{\boldsymbol{u}}$ with $k > h+1$ do not occur with a non-vanishing coefficient.

From

$$\underset{h}{\boldsymbol{u}}\underset{k}{\boldsymbol{u}} = \begin{cases} 1, & \text{if } h = k \\ 0, & \text{if } h \neq k \end{cases}$$

we deduce

$$\underset{k}{\boldsymbol{u}}\,\mathrm{d}_s\underset{h}{\boldsymbol{u}} + \underset{h}{\boldsymbol{u}}\,\mathrm{d}_s\underset{k}{\boldsymbol{u}} = 0. \tag{6.2–4}$$

It follows from (6.2–3), that $\underset{kh}{\alpha} = \underset{k}{\boldsymbol{u}}\,\mathrm{d}_s\underset{h}{\boldsymbol{u}}$. Hence (6.2–4) is equivalent to

$$\underset{hk}{\alpha}+\underset{kh}{\alpha} = 0, \tag{6.2–5}$$

that is to say, the matrix

$$[\underset{hk}{\alpha}] \tag{6.2–6}$$

is skew symmetric. Thus also $\underset{kh}{\alpha} = 0$, whenever $k < h-1$. We now may simplify (6.2–3) by introducing the numbers

$$\underset{h}{\kappa} = \underset{h\,h-1}{\alpha}\,, \qquad h = 1, \ldots, \mathrm{r}.$$

Then the equations (6.2–3) (save the first and the last), appear as

$$\boxed{\mathrm{d}_s \underset{h}{\boldsymbol{u}} = -\underset{h-1}{\boldsymbol{u}}\,\underset{h}{\kappa} + \underset{h+1}{\boldsymbol{u}}\,\underset{h+1}{\kappa}\,,} \quad h = 1, \ldots, \mathrm{r}-1. \tag{6.2–7}$$

Additionally we have

$$\boxed{\mathrm{d}_s \underset{0}{\boldsymbol{u}} = \underset{1}{\boldsymbol{u}}\,\underset{1}{\kappa},} \tag{6.2–8}$$

being essentially the same as formula (6.1–14).

The expressions $\kappa(s)$ are determined except for sign. We agree to take them positive. Then the vectors $\underset{h}{\boldsymbol{u}}$ are also uniquely determined. Comparing (6.2–8) with (6.1–14) we see that $\underset{1}{\kappa}$ is the curvature. The functions $\underset{2}{\kappa}, \underset{3}{\kappa}, \ldots, \underset{\mathrm{r}}{\kappa}$ are called the *torsions* of the curve.

Finally we wish to make the following remark. Since the curve is included in a finite dimensional space of dimension $r+1$ the derivative $\mathrm{d}_s\underset{\mathrm{r}}{\boldsymbol{u}}$ also belongs to this space, which means that the equation corresponding to $h = \mathrm{r}$ in (6.2–7) takes the form

$$\boxed{\mathrm{d}_s \underset{\mathrm{r}}{\boldsymbol{u}} = -\underset{\mathrm{r}-1}{\boldsymbol{u}}\,\underset{\mathrm{r}}{\kappa}.} \tag{6.2–9}$$

The relations (6.2–7), (6.2–8) and (6.2–9) with be referred to as *Frenet's equations.*

6.2.2 – Explicit expressions for the curvature and the torsions

It is not difficult to express the curvature and the torsions in terms of the successive derivatives (6.1–3) of the function $\boldsymbol{x}(q)$. First we suppose that the parameter is the arc length. Since $\boldsymbol{u}$ is included in the space spanned by

$d_s x, \ldots, d_s^{h+1} x$, we may write

$$\underset{h}{u} = d_s x \underset{h1}{\beta} + d_s^2 x \underset{h2}{\beta} + \ldots + d_s^{h+1} x \underset{hh+1}{\beta} . \qquad (6.2\text{–}10)$$

Since $\underset{h}{u}$ is orthogonal to $d_s x, \ldots, d_s^2 x$, we have

$$1 = \underset{h}{u}\, d_s^{h+1} x \underset{hh+1}{\beta} . \qquad (6.2\text{–}11)$$

From Frenet's equation we may infer that

$$\underset{h}{\kappa} = - \underset{h-1}{u}\, d_s \underset{h}{u} = \underset{h}{u}\, d_s \underset{h-1}{u} \qquad (6.2\text{–}12)$$

and, when differentiating both members of (6.2–10), (replacing h by $h-1$) it easily follows that

$$\underset{h}{\kappa} = \underset{h}{u}\, d_s^{h+1} x \underset{h-1h}{\beta} , \qquad (6.2\text{–}13)$$

whence by virtue of (6.2–11)

$$\underset{h}{\kappa} = \underset{h-1h}{\beta} \,/\, \underset{hh+1}{\beta} . \qquad (6.2\text{–}14)$$

Let

$$\operatorname{gr}(d_s x, \ldots, d_s^h x) \qquad (6.2\text{–}15)$$

denote the gramian of the vectors $d_s x, \ldots, d_s^h x$. It follows from (6.2–10), since the gramian of $\underset{0}{u}, \ldots, \underset{h}{u}$ is unity, that

$$1 = \underset{01}{\beta^2}\ \underset{12}{\beta^2} \ldots \underset{hh+1}{\beta^2} \operatorname{gr}(d_s x, \ldots, d_s^{h+1} x) \qquad (6.2\text{–}16)$$

and an easy computation yields from (6.2–14) and (6.2–16)

$$\underset{h}{\kappa^2} = \frac{\operatorname{gr}(d_s x, \ldots, d_s^{h-1} x) \operatorname{gr}(d_s x, \ldots, d_s^{h+1} x)}{\operatorname{gr}^2(d_s x, \ldots, d_s^h x)}, \quad h > 1. \qquad (6.2\text{–}17)$$

Additionally we have

$$\underset{01}{\beta} = \frac{1}{d_s x\, d_s x} .$$

Hence

$$\underset{1}{\kappa} = \frac{\operatorname{gr}(d_s x, d_s^2 x)}{\operatorname{gr}^2(d_s x)}, \qquad (6.2\text{–}18)$$

the same expression as (6.2–17), when we agree to take the gramian of an empty set as being equal to one.

With reference to relations of the type (6.1–4) we easily find that

$$\operatorname{gr}(d_s x, \ldots, d_s^h x) = \operatorname{gr}(d_q x, \ldots, d_q^h x)(d_s q)^{h(h+1)}, \qquad (6.2\text{–}19)$$

in particular

$$1 = \mathrm{d}_s x\, \mathrm{d}_s x = \mathrm{d}_q x\, \mathrm{d}_q x (\mathrm{d}_s q)^2 = \mathrm{gr}(\mathrm{d}_q x)(\mathrm{d}_s q)^2. \qquad (6.2\text{–}20)$$

When the curve is referred to an arbitrary parameter q we, therefore, have

$$\boxed{\underset{h}{\kappa}^2 = \frac{\mathrm{gr}(\mathrm{d}_q x, \ldots, \mathrm{d}_q^{h-1} x)\mathrm{gr}(\mathrm{d}_q x, \ldots, \mathrm{d}_q^{h+1} x)}{\mathrm{gr}(\mathrm{d}_q x)\mathrm{gr}^2(\mathrm{d}_q x, \ldots, \mathrm{d}_q^{h} x)},} \quad h = 1, \ldots, r. \qquad (6.2\text{–}21)$$

This formula includes the special case

$$\underset{1}{\kappa}^2 = \frac{\mathrm{gr}(\mathrm{d}_q x, \mathrm{d}_q^2 x)}{\mathrm{gr}^3(\mathrm{d}_q x)}. \qquad (6.2\text{–}22)$$

By hypothesis the denominator in (6.2–21) does not vanish identically, that is to say, the curve is not included in a space of dimension less than $r+1$.

6.2.3 – A THEOREM ABOUT THE VARIATION OF THE ARC LENGTH

Besides a given curve $x(s)$ we consider a second curve

$$x(s; \varepsilon) = x(s) + y(s)\varepsilon, \qquad (6.2\text{–}23)$$

where ε is a constant. This curve is referred to the arc length of the given curve, but s need not be the arc length of (6.2–23). The length of the arc between the points $s = a$ and $s = b$ is equal to

$$s(\varepsilon) = \int_a^b \sqrt{(\mathrm{d}_s x + \mathrm{d}_s y\varepsilon)(\mathrm{d}_s x + \mathrm{d}_s y\varepsilon)}\,\mathrm{d}s. \qquad (6.2\text{–}24)$$

By a *variation* of the curve we understand the additional term in (6.2–23). By the *first variation of the arc length* is understood the expression

$$\delta s = \mathrm{d}_\varepsilon s(\varepsilon)\big|_{\varepsilon=0}\, \varepsilon. \qquad (6.2\text{–}25)$$

An easy computation yields from (6.2–24)

$$\delta s = \varepsilon \int_a^b \mathrm{d}_s x\, \mathrm{d}_s y\, \mathrm{d}s = \varepsilon \int_a^b \underset{0}{u}\, \mathrm{d}_s y\, \mathrm{d}s. \qquad (6.2\text{–}26)$$

Integrating by parts yields

$$\delta s = \varepsilon \underset{0}{u} y \Big|_a^b - \varepsilon \int_a^b \underset{1}{u} y \underset{1}{\kappa}\, \mathrm{d}s$$

or, by putting

$$\delta x = y\varepsilon,$$

$$\delta s = \underset{0}{\boldsymbol{u}}\,\delta \boldsymbol{x}\Big|_a^b - \int_a^b \underset{1}{\boldsymbol{u}}\,\delta \boldsymbol{x} \underset{1}{\kappa} \mathrm{d}s. \tag{6.2–27}$$

This result implies the following theorem:

The first variation of the arc vanishes when variation of the curve takes place in a direction orthogonal to the osculating plane everywhere along the curve throughout the range $a \leqq s \leqq b$.

This follows since $\underset{0}{\boldsymbol{u}}\delta\boldsymbol{x} = \underset{1}{\boldsymbol{u}}\delta\boldsymbol{x} = 0$ for $a \leqq s \leqq b$.

6.2.4 – Hyperspherical curves

Suppose that $\boldsymbol{x}(s)$ represents a curve which is included in an $(r+1)$-dimensional space (and not in a space of lower dimension) such that all its points are at a constant distance from a fixed point, the *centre,* with coordinate vector $\boldsymbol{a}$, i.e.,

$$(\boldsymbol{x}(s)-\boldsymbol{a})(\boldsymbol{x}(s)-\boldsymbol{a}) = \text{const.} \tag{6.2–28}$$

A curve having this property is called *hyperspherical;* it is a curve on the *hyperspere* (6.2–28). We wish to establish the theorem:

A necessary and sufficient condition for a curve in an $(r+1)$-dimensional space to be hyperspherical is the existence of r functions

$$\underset{1}{\xi}, \ldots, \underset{r}{\xi}$$

which satisfy the conditions

$$\begin{gathered} 1-\underset{1}{\xi}\,\underset{1}{\kappa} = 0, \\ \mathrm{d}_s\underset{1}{\xi} = \underset{2}{\xi}\,\underset{2}{\kappa}, \qquad \mathrm{d}_s\underset{r}{\xi} = -\underset{r-1}{\xi}\,\underset{r}{\kappa}, \\ \mathrm{d}_s\underset{h}{\xi} = -\underset{h-1}{\xi}\,\underset{h}{\kappa} + \underset{h+1}{\xi}\,\underset{h+1}{\kappa}, \qquad h = 2, \ldots, r-1. \end{gathered} \tag{6.2–29}$$

By a *normal r-space* at a given point of the curve we understand the space spanned by the vectors $\underset{1}{\boldsymbol{u}}, \ldots, \underset{r}{\boldsymbol{u}}$ at this point. It is easy to show that the normal r-space of a hyperspherical curve passes through the centre of the hypersphere on which the curve lies. In fact, by differentiating (6.2–28) we get

$$(\boldsymbol{x}(s)-\boldsymbol{a})\mathrm{d}_s\boldsymbol{x} = 0, \tag{6.2–30}$$

expressing the fact that the vector $\boldsymbol{x}(s)-\boldsymbol{a}$ is orthogonal to the tangent vector $\underset{0}{\boldsymbol{u}} = \mathrm{d}_s\boldsymbol{x}$.

As a consequence there are functions $\xi_1, \ldots, \xi_r$ such that

$$\boldsymbol{a} = \boldsymbol{x}(s) + \sum_{h=1}^{r} \boldsymbol{u}_h \xi_h. \tag{6.2–31}$$

By differentiating both sides of this equation we get

$$\begin{aligned} \boldsymbol{o} &= \boldsymbol{u}_0 + \sum_{h=1}^{r-1} (-\boldsymbol{u}_{h-1} \kappa_h + \boldsymbol{u}_{h+1} \kappa_{h+1}) \xi_h - \boldsymbol{u}_{r-1} \kappa_r \xi_r + \sum_{h=1}^{r} \boldsymbol{u}_h \mathrm{d}_s \xi_h \\ &= \boldsymbol{u}_0 (1 - \xi_1 \kappa_1) + \boldsymbol{u}_1 (\mathrm{d}_s \xi_1 - \xi_2 \kappa_2) + \sum_{h=2}^{r-1} \boldsymbol{u}_h (\mathrm{d}_s \xi_h + \xi_{h-1} \kappa_h - \xi_{h+1} \kappa_{h+1}) + \boldsymbol{u}_r (\mathrm{d}_s \xi_r + \xi_{r-1} \kappa_r). \end{aligned} \tag{6.2–32}$$

Since at every point of the curve the vectors $\boldsymbol{u}_0, \ldots, \boldsymbol{u}_r$ are linearly independent, we obtain the desired equations (6.2–29).

Conversely, if a set of ξ's can be found, such that the conditions (6.2–29) have been fulfilled, then the derivative of the function represented by the right hand side of (6.2–31), being equal to the right hand side of (6.2–32), vanishes identically. As a consequence this vector function is a constant vector $\boldsymbol{a}$. In addition we may infer from (6.2–29) that

$$\sum_{h=1}^{r} \xi_h \mathrm{d}_s \xi_h = 0,$$

i.e.,

$$\sum_{h=1}^{r} \xi_h^2 = \text{constant}.$$

This equation expresses the fact that the norm of the vector $\boldsymbol{a} - \boldsymbol{x}(s)$ is constant along the curve. This completes the proof of the assertion.

It is not difficult to eliminate the auxiliary functions $\xi_1, \ldots, \xi_r$ in order to obtain a relation between the curvatures. The result is, however, rather intricate. In the case $r = 2$ the last equations (6.2–29) are void and we obtain

$$\kappa_2/\kappa_1 + \mathrm{d}_s(\mathrm{d}_s(1/\kappa_1)/\kappa_2) = 0,$$

a relation between the principal curvature and the torsion of a spherical curve in the three dimensional space.

6.2.5 – Helices

A curve $\boldsymbol{x}(s)$ in a space whose dimension is an *odd* number $r+1$ is called a *helix* if the vectors $\boldsymbol{u}_0, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{r-2}$ make constant angles with a fixed unit vector $\boldsymbol{u}$. It is our aim to prove:

If the vectors

$$\mathbf{u}_0, \mathbf{u}_2, \ldots, \mathbf{u}_{r-2}$$

make constant angles with a given unit vector $\mathbf{u}$ *then the vectors*

$$\mathbf{u}_1, \mathbf{u}_3, \ldots, \mathbf{u}_{r-1}$$

are orthogonal to $\mathbf{u}$ *and the ratios*

$$\kappa_1/\kappa_2, \kappa_3/\kappa_4, \ldots, \kappa_{r-1}/\kappa_r \tag{6.2–33}$$

of the curvatures are constant along the curve. Moreover the vector $\mathbf{u}_r$ *makes a constant angle with* $\mathbf{u}$.

Denote the angles mentioned in the theorem by

$$\omega_0, \omega_2, \ldots, \omega_r,$$

that is,

$$\mathbf{u}\mathbf{u}_0 = \cos \omega_0, \ \mathbf{u}\mathbf{u}_2 = \cos \omega_2, \ldots, \mathbf{u}\mathbf{u}_r = \cos \omega_r. \tag{6.2–34}$$

By hypothesis all these angles are constant, except perhaps the last one. By differentiating the equations (6.2–34) except the last equation, we obtain on applying Frenet's formulas:

$$\mathbf{u}\mathbf{u}_1\kappa_1 = 0, \quad -\mathbf{u}\mathbf{u}_1\kappa_1 + \mathbf{u}\mathbf{u}_3\kappa_3 = 0, \ldots, -\mathbf{u}\mathbf{u}_{r-3}\kappa_{r-2} + \mathbf{u}\mathbf{u}_{r-1}\kappa_{r-1} = 0. \tag{6.2–35}$$

Since none of the curvatures $\kappa_1, \ldots, \kappa_{r-2}$ vanishes identically it follows that

$$\mathbf{u}\mathbf{u}_1 = \mathbf{u}\mathbf{u}_3 = \ldots = \mathbf{u}\mathbf{u}_{r-1} = 0. \tag{6.2–36}$$

Again by differentiating the equations (6.2–36) we have

$$(-\mathbf{u}_0\kappa_1 + \mathbf{u}_2\kappa_2)\mathbf{u} = (-\mathbf{u}_2\kappa_3 + \mathbf{u}_4\kappa_4)\mathbf{u} = \ldots = (-\mathbf{u}_{r-2}\kappa_{r-1} + \mathbf{u}_r\kappa_r)\mathbf{u} = 0$$

or, in view of (6.2–34),

$$\kappa_1 \cos \omega_0 = \kappa_2 \cos \omega_2, \ \kappa_3 \cos \omega_2 = \kappa_4 \cos \omega_4, \ldots, \kappa_{r-1} \cos \omega_{r-2} = \kappa_r \cos \omega_r.$$

Next we observe that

$$d_s(\mathbf{u}\mathbf{u}_r) = -\mathbf{u}\mathbf{u}_{r-1}\kappa_r = 0.$$

Hence ω_r also is constant and the proof is complete.

Conversely we have the theorem:

If the ratios (6.2–33) *are constant along a curve* $\mathbf{x}(s)$ *in an* $(r+1)$*-dimensional space, where* $r+1$ *is odd, then the curve is a helix.*

By hypothesis we can find constants $\gamma_0, \gamma_2, \ldots, \gamma_r$ such that

$$\kappa_1\gamma_0 = \kappa_2\gamma_2,\ \kappa_3\gamma_2 = \kappa_4\gamma_4,\ \ldots,\ \kappa_{r-1}\gamma_{r-2} = \kappa_r\gamma_r \tag{6.2–37}$$

and

$$\gamma_0^2+\gamma_2^2+\ldots+\gamma_r^2 = 1. \tag{6.2–38}$$

Now we consider the vector function

$$\boldsymbol{u}(s) = \boldsymbol{u}_0\gamma_0+\boldsymbol{u}_2\gamma_2+\ldots+\boldsymbol{u}_r\gamma_r. \tag{6.2–39}$$

Differentiation yields

$$\begin{aligned} \mathrm{d}_s\boldsymbol{u} &= \boldsymbol{u}_1\kappa_1\gamma_0+(-\boldsymbol{u}_1\kappa_2+\boldsymbol{u}_3\kappa_3)\gamma_2+\ldots+(-\boldsymbol{u}_{r-1}\kappa_r)\gamma_r \\ &= \boldsymbol{u}_1(\kappa_1\gamma_0-\kappa_2\gamma_2)+\boldsymbol{u}_3(\kappa_3\gamma_2-\kappa_4\gamma_4)+\ldots+\boldsymbol{u}_r(\kappa_{r-1}\gamma_{r-2}-\kappa_r\gamma_r) = \boldsymbol{o}. \end{aligned}$$

Hence $\boldsymbol{u}$ is constant along the curve and from (6.2–38) we may even infer that $\boldsymbol{u}$ is a unit vector. Finally it is seen that the angles between $\boldsymbol{u}$ and $\boldsymbol{u}_0, \ldots, \boldsymbol{u}_r$ are determined by

$$\cos\omega_0 = \gamma_0,\ \cos\omega_2 = \gamma_2,\ \ldots,\ \cos\omega_r = \gamma_r,$$

and are, therefore, constant.

6.2.6 – Orthogonal trajectories of the normal $(r-1)$-spaces of a curve

The space spanned by the normals $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{r-1}$ at a given point of a curve in an $(r+1)$-dimensional space is called the *normal* $(r-1)$-*space* of the curve at this point. We intend to study a curve related to a given curve $\boldsymbol{x}(s)$ by

$$\boldsymbol{y}(s) = \boldsymbol{x}(s)+\sum_{h=1}^{r-1}\boldsymbol{u}_h\xi_h, \tag{6.2–40}$$

where $\xi_1, \ldots, \xi_{r-1}$ denote functions of the arc s. The tangent line of this curve at the point s is determined by the vector

$$\mathrm{d}_s\boldsymbol{y} = \boldsymbol{u}_0(1-\xi_1\kappa_1)+\boldsymbol{u}_1(\mathrm{d}_s\xi_1+\xi_2\kappa_2)+\sum_{h=2}^{r-2}\boldsymbol{u}_h(\mathrm{d}_s\xi_h-\xi_{h-1}\kappa_{h-1}+\xi_{h+1}\kappa_{h+1})+\boldsymbol{u}_r\xi_{r-1}\kappa_r, \tag{6.2–41}$$

Now we impose on this curve the further condition that this tangent is orthogonal to the normal $(r-1)$-space of the original curve at the same point s.

Then the curve $\boldsymbol{y}(s)$ is called an *orthogonal trajectory* of the normal $(r-1)$-spaces of the curve $x(s)$. We shall prove:

If the curve (6.2–40) *is an orthogonal trajectory of the normal* $(r-1)$*-spaces of the curve* $\boldsymbol{x}(s)$*, then the distance of corresponding points of the curves* $\boldsymbol{y}(s)$ *and* $\boldsymbol{x}(s)$ *is constant along the curve.*

The vector $d_s\boldsymbol{y}$ is orthogonal to $\underset{1}{\boldsymbol{u}}, \ldots, \underset{r-1}{\boldsymbol{u}}$ when the following conditions have been satisfied:

$$\begin{aligned} d_s\underset{1}{\xi}+\underset{2}{\xi}\underset{2}{\kappa} &= 0, \\ d_s\underset{h}{\xi}-\underset{h-1}{\xi}\underset{h}{\kappa}+\underset{h+1}{\xi}\underset{h+1}{\kappa} &= 0, \qquad h = 2, \ldots, r-1. \end{aligned} \tag{6.2–42}$$

From these equations we easily deduce

$$\sum_{h=1}^{r-1} \underset{h}{\xi} d_s \underset{h}{\xi} = 0,$$

whence

$$\sum_{h=1}^{r-1} \underset{h}{\xi}^2 = \text{constant}. \tag{6.2–43}$$

This means, however, that the norm of the vector $\boldsymbol{y}-\boldsymbol{x}$ is constant along the given curve $\boldsymbol{x}(s)$.

In the remainder of this section we always suppose that $\boldsymbol{y}(s)$ is an orthogonal trajectory of the normal $(r-1)$-spaces of a curve $\boldsymbol{x}(s)$. In accordance with (6.2–42) the equation (6.2–41) appears in the simple form

$$d_s\boldsymbol{y} = \underset{0}{\boldsymbol{u}}(1-\underset{1}{\xi}\underset{1}{\kappa})+\underset{r}{\boldsymbol{u}}\underset{r-1}{\xi}\underset{r}{\kappa}. \tag{6.2–44}$$

Let $\underset{0}{\boldsymbol{v}}$ denote a unit vector parallel to $d_s\boldsymbol{y}$. It is clear that we may write

$$\underset{0}{\boldsymbol{v}} = \underset{0}{\boldsymbol{u}} \cos \omega+\underset{r}{\boldsymbol{u}} \sin \omega. \tag{6.2–45}$$

If $\sin \omega = 0$ the curves $\boldsymbol{x}(s)$ and $\boldsymbol{y}(s)$ are said to be *parallel*. We exclude this case from further considerations.

Next we impose on $\boldsymbol{y}(s)$ the additional condition that its first normal vector $\underset{1}{\boldsymbol{v}}$ is included in the corresponding normal $(r-1)$-space of $\boldsymbol{x}(s)$. This leads to a result that may be stated as:

When the first normal of an orthogonal trajectory of the normal $(r-1)$*-spaces of a curve* $\boldsymbol{x}(s)$ *is included in the corresponding normal* $(r-1)$*-space then the angle between the tangent vectors at corresponding points is constant.*

The normal $\underset{1}{\boldsymbol{v}}$ is parallel to $\mathrm{d}_s \underset{0}{\boldsymbol{v}}$. Now

$$\mathrm{d}_s \underset{0}{\boldsymbol{v}} = \underset{1}{\boldsymbol{u}} \underset{1}{\kappa} \cos\omega + \underset{0}{\boldsymbol{u}}\, \mathrm{d}_s \cos\omega - \underset{r-1}{\boldsymbol{u}} \underset{r}{\kappa} \sin\omega + \underset{r}{\boldsymbol{u}}\, \mathrm{d}_s \sin\omega.$$

Since, by hypothesis $\mathrm{d}_s \underset{0}{\boldsymbol{v}}$ is linearly dependent on $\underset{1}{\boldsymbol{u}}, \ldots, \underset{r-1}{\boldsymbol{u}}$, we may conclude that

$$\mathrm{d}_s \sin\omega = 0,$$

i.e., ω is constant, ω being the angle between $\underset{0}{\boldsymbol{u}}$ and $\underset{0}{\boldsymbol{v}}$.

Comparing (6.2–44) and (6.2–45) we conclude that

$$\begin{vmatrix} 1-\underset{1}{\xi}\underset{1}{\kappa} & \underset{r-1}{\xi}\underset{r}{\kappa} \\ \cos\omega & \sin\omega \end{vmatrix} = 0,$$

i.e.,

$$1 = \underset{1}{\xi}\underset{1}{\kappa} + \underset{r-1}{\xi}\underset{r}{\kappa} \operatorname{ctn}\omega. \tag{6.2–46}$$

An interesting case occurs when $r = 2$. Then $\boldsymbol{x}(s)$ and $\boldsymbol{y}(s)$ constitute a pair of curves in a three dimensional space being related in such a way that the line connecting two corresponding points is normal to both curves. Such a pair is called a *Bertrand pair of curves*. From (6.2–43) follows that $\underset{1}{\xi}$ is constant. Hence, since $\underset{1}{\xi} = \underset{r-1}{\xi}$, (6.2–46) reads

$$1 = \xi(\underset{1}{\kappa} + \underset{2}{\kappa} \operatorname{ctn}\omega). \tag{6.2–47}$$

Thus we see that there exists a linear relation between the principal curvature and and the torsion of the curve $\boldsymbol{x}(s)$. It is not difficult to show that the same statement is true for $\boldsymbol{y}(s)$.

6.3 – Involutes and evolutes

6.3.1 – Involutes

The points of the tangent lines of a curve $x(s)$ constitute a surface, the *tangent surface of the curve*. It is a ruled surface, whose generators are the tangent lines of the given curve. This curve forms a sharp line on the surface, the *edge of regression*.

We are interested in the curves on the tangent surface, which meet the generators at right angles. They are called the *involutes* of the curve.

Any curve on the tangent surface is characterized by

$$\boldsymbol{y}(s) = \boldsymbol{x}(s) + \underset{0}{\boldsymbol{u}}\xi(s), \tag{6.3–1}$$

where $\xi(s)$ is a function of s. It may be determined when we impose on it the condition, that the vectors $d_s \boldsymbol{y}$ and $\underset{0}{\boldsymbol{u}}$ are orthogonal for all values of s. From (6.3–1) we have

$$d_s \boldsymbol{y} = \underset{0}{\boldsymbol{u}}(1+d_s\xi)+\underset{1}{\boldsymbol{u}}\underset{1}{\kappa}\xi \tag{6.3–2}$$

and from $\underset{0}{\boldsymbol{u}} d_s \boldsymbol{y} = 0$ we get

$$1+d_s\xi = 0.$$

Hence

$$\xi = c-s,$$

c being an arbitrary constant. Thus we have proved:

A curve $\boldsymbol{x}(s)$ possesses a one parameter family of involutes represented by

$$\boldsymbol{y}(s) = \boldsymbol{x}(s)+\underset{0}{\boldsymbol{u}}(s)(c-s), \tag{6.3–3}$$

where c is an arbitrary constant and s is the arc length of the original curve.

It should be noticed that two curves of the family cut segments of equal length on the generators of the tangent surface. They are *equidistant* on this surface.

6.3.2 – Involutes of higher order

It is an easy matter to generalize the notion of involute in the following way. Consider a given curve $\boldsymbol{x}(s)$ and the curve

$$\boldsymbol{y}(s) = \boldsymbol{x}(s)+\underset{0}{\boldsymbol{u}}\underset{0}{\xi}+\ldots+\underset{n-1}{\boldsymbol{u}}\underset{n-1}{\xi}, \tag{6.3–4}$$

where $\underset{0}{\xi},\ldots,\underset{n-1}{\xi}$, $n \leqq r$, are functions of s. We ask for those curves (6.3–4) which are orthogonal trajectories of the system of the n-dimensional osculating spaces of $\boldsymbol{x}(s)$. Curves with this property are called *involutes of order n* of the given curve.

By virtue of Frenet's formulas we evidently have

$$\begin{aligned} d_s \boldsymbol{y} = {} & \underset{0}{\boldsymbol{u}}(1+d_s\underset{0}{\xi}-\underset{1}{\xi}\underset{1}{\kappa})+\sum_{h=2}^{n-2}\underset{h}{\boldsymbol{u}}(d_s\underset{h}{\xi}+\underset{h-1}{\xi}\underset{h}{\kappa}-\underset{h+1}{\xi}\underset{h+1}{\kappa}) \\ & +\underset{n-1}{\boldsymbol{u}}(d_s\underset{n-1}{\xi}+\underset{n-2}{\xi}\underset{n-1}{\kappa})+\underset{n}{\boldsymbol{u}}\underset{n-1}{\xi}\underset{n}{\kappa}, \end{aligned} \tag{6.3–5}$$

Since by hypothesis, $d_s \boldsymbol{y}$, is orthogonal to $\underset{0}{\boldsymbol{u}},\ldots,\underset{n-1}{\boldsymbol{u}}$, we obtain the system of differential equations

$$\begin{aligned} 0 &= 1+\mathrm{d}_s\underset{0}{\xi}-\underset{1}{\xi}\,\underset{1}{\kappa}, \\ 0 &= \mathrm{d}_s\underset{h}{\xi}+\underset{h-1}{\xi}\,\underset{h}{\kappa}-\underset{h+1}{\xi}\,\underset{h+1}{\kappa}, \qquad h = 1,\ldots,\mathrm{n}-2, \\ 0 &= \mathrm{d}_s\underset{n-1}{\xi}+\underset{n-2}{\xi}\,\underset{n-1}{\kappa}, \end{aligned} \tag{6.3–6}$$

which admits of a uniquely determined set of solutions $\underset{0}{\xi},\ldots,\underset{\mathrm{n}-1}{\xi}$, having prescribed initial values at a point $s = a$ of the given curve. In the case $n = 1$ these equations reduce to $1+\mathrm{d}_s\underset{0}{\xi} = 0$.

6.3.3 – Evolutes

We wish to discuss the converse problem, dealt with in section 6.3.1, viz., to find a curve $\boldsymbol{y}(s)$ such that a given curve $\boldsymbol{x}(s)$ is one of its involutes (of the first order). A curve of this kind is called an *evolute* of the given curve.

It is not always true that a curve possesses an evolute. A counter-example is the circle.

Next we shall prove:

A necessary and sufficient condition for a curve $\boldsymbol{x}(s)$ to posses an evolute is that there is a unit vector function $\boldsymbol{u}(s)$ and a function $\zeta(s)$, such that

$$\underset{0}{\boldsymbol{u}}+\mathrm{d}_s\boldsymbol{u}\,\zeta = \boldsymbol{o}. \tag{6.3–7}$$

Let $\boldsymbol{y}(s)$ denote an involute of the curve $\boldsymbol{x}(s)$. Then $\boldsymbol{y}(s)-\boldsymbol{x}(s)$ is orthogonal to $\underset{0}{\boldsymbol{u}}$ and $\mathrm{d}_s\boldsymbol{y}$ linearly dependent on $\boldsymbol{y}-\boldsymbol{x}$. It follows that $\mathrm{d}_s\boldsymbol{y}$ is also orthogonal to $\underset{0}{\boldsymbol{u}}$ and that $\boldsymbol{y}$ and $\boldsymbol{x}$ do not coincide. As a consequence we can find a unit vector $\boldsymbol{u}$ such that

$$\boldsymbol{y}-\boldsymbol{x} = \boldsymbol{u}\zeta \tag{6.3–8}$$

and

$$\mathrm{d}_s\boldsymbol{y} = \boldsymbol{u}\xi. \tag{6.3–9}$$

By differentiating both members of (6.3–8) we get

$$\boldsymbol{u}\xi-\underset{0}{\boldsymbol{u}} = \mathrm{d}_s\boldsymbol{u}\,\zeta+\boldsymbol{u}\,\mathrm{d}_s\zeta$$

or

$$\boldsymbol{u}(\xi-\mathrm{d}_s\zeta) = \underset{0}{\boldsymbol{u}}+\mathrm{d}_s\boldsymbol{u}\,\zeta. \tag{6.3–10}$$

Multiplying both members of this equation by $\boldsymbol{u}$ we get $\xi-\mathrm{d}_s\zeta = 0$ and (6.3–7) follows.

Assume, conversely, that a relation of the type (6.3–7) holds for a suitably chosen unit vector $\boldsymbol{u}(s)$ and a function $\zeta(s)$. Define $\boldsymbol{y}(s)$ by

$$\boldsymbol{y}(s) = \boldsymbol{x}(s)+\boldsymbol{u}(s)\zeta(s). \tag{6.3–11}$$

Then

$$\mathrm{d}_s\boldsymbol{y} = \underset{0}{\boldsymbol{u}}+\mathrm{d}_s\boldsymbol{u}\,\zeta+\boldsymbol{u}\,\mathrm{d}_s\zeta = \boldsymbol{u}\,\mathrm{d}_s\zeta,$$

hence $\boldsymbol{y}-\boldsymbol{x}$ and $\mathrm{d}_s\boldsymbol{y}$ are linearly dependent. Also

$$\underset{0}{\boldsymbol{u}}\,\mathrm{d}_s\boldsymbol{y} = \underset{0}{\boldsymbol{u}}\boldsymbol{u}\,\mathrm{d}_s\zeta = -\boldsymbol{u}\,\mathrm{d}_s\boldsymbol{u}\,\zeta\,\mathrm{d}_s\zeta = 0,$$

for $\boldsymbol{u}$ is a unit vector. It follows that $\mathrm{d}_s\boldsymbol{y}$ is orthogonal to $\mathrm{d}_s\boldsymbol{x}$ and this concludes the proof of the theorem.

By the *centre of curvature* at a point s of a curve $\boldsymbol{x}(s)$ we mean a point whose coordinate vector is

$$\boldsymbol{x}(s)+\underset{1}{\boldsymbol{u}}(s)\rho \tag{6.3–12}$$

where $\rho = \kappa^{-1}$ denotes the *radius of curvature* at the point under consideration. We shall prove:

The projection of a point of an evolute onto the first normal at the corresponding point of the curve is the centre of curvature at that point.

The projection referred to in the theorem is given by the vector

$$\boldsymbol{x}+\underset{1}{\boldsymbol{u}}(\underset{1}{\boldsymbol{u}}\boldsymbol{z}), \tag{6.3–13}$$

where

$$\boldsymbol{z} = \boldsymbol{y}-\boldsymbol{x}. \tag{6.3–14}$$

Since $\boldsymbol{z}$ is orthogonal to $\underset{0}{\boldsymbol{u}}$ we have

$$\underset{0}{\boldsymbol{u}}\,\mathrm{d}_s\boldsymbol{z}+\underset{1}{\boldsymbol{u}}\underset{1}{\boldsymbol{z}}\kappa = 0. \tag{6.3–15}$$

In addition we have

$$\boldsymbol{z} = \boldsymbol{u}\zeta,$$

where $\boldsymbol{u}$ and ζ are the quantities mentioned in the above theorem. In view of (6.3–7) we have

$$\mathrm{d}_s\boldsymbol{z} = \mathrm{d}_s\boldsymbol{u}\,\zeta+\boldsymbol{u}\,\mathrm{d}_s\zeta = -\underset{0}{\boldsymbol{u}}+\boldsymbol{u}\,\mathrm{d}_s\zeta.$$

Hence

$$\underset{0}{\boldsymbol{u}}\,\mathrm{d}_s\boldsymbol{z} = -1$$

and we find that in (6.3–15) the product $\underset{1}{\boldsymbol{u}}\boldsymbol{z}$ is equal to $\kappa^{-1} = \rho$. Inserting this into (6.3–13) we obtain the expression (6.3–12), as desired.

Finally we wish to prove the theorem:

The lines from a point of a curve which are tangent to two of its involutes meet at a constant angle along the curve.

By the first theorem of this section, these evolutes are characterized by vector functions $'\boldsymbol{u}$, $''\boldsymbol{u}$ and functions $'\zeta$, $''\zeta$, such that

$$\underset{0}{\boldsymbol{u}}+\mathrm{d}_s{}'\boldsymbol{u}'\zeta = \underset{0}{\boldsymbol{u}}+\mathrm{d}_s{}''\boldsymbol{u}''\zeta = \boldsymbol{o}.$$

Since $'\boldsymbol{u}$ and $''\boldsymbol{u}$ are orthogonal to $\underset{0}{\boldsymbol{u}}$, we evidently have

$$'\boldsymbol{u}\mathrm{d}_s{}''\boldsymbol{u} = {}''\boldsymbol{u}\mathrm{d}_s{}'\boldsymbol{u} = 0,$$

whence

$$\mathrm{d}_s('\boldsymbol{u}\,''\boldsymbol{u}) = 0.$$

This completes the proof.

6.4 – Developables

6.4.1 – General remarks

A one parameter system of *hyperplanes* in an $(r+1)$-dimensional space can be represented by

$$\boldsymbol{y}\boldsymbol{p}(q) = \eta(q), \tag{6.4–1}$$

where $\boldsymbol{p}(q)$ and $\eta(q)$ are functions of a parameter q. Let q denote a given number and $q+h$ a number near it. It determines a second hyperplane

$$\boldsymbol{y}\boldsymbol{p}(q+h) = \eta(q+h). \tag{6.4–2}$$

The hyperplanes (6.4–1) and (6.4–2) have the same intersection as the hyperplanes (6.4–1) and

$$\boldsymbol{y}(\boldsymbol{p}(q+h)-\boldsymbol{p}(q))\frac{1}{h} = \frac{\eta(q+h)-\eta(q)}{h}. \tag{6.4–3}$$

When we let $h \to 0$ we obtain the hyperplane

$$\boldsymbol{y}\mathrm{d}_q\boldsymbol{p} = \mathrm{d}_q\eta. \tag{6.4–4}$$

Performing the same operation on (6.5–4) we obtain a hyperplane

$$\boldsymbol{y}\mathrm{d}_q^2\boldsymbol{p} = \mathrm{d}_q^2\eta \tag{6.4–5}$$

and so on. The intersection of the hyperplanes

$$yp = \eta,\ y\mathrm{d}_q p = \mathrm{d}_q \eta,\ \ldots,\ y\mathrm{d}_q^{\ m} p = \mathrm{d}_q^{\ m} \eta \qquad (6.4\text{–}6)$$

is, in general, a linear space of dimension $r-m$. It is called the *characteristic* of dimension $r-m$ of the system. For $m = r$ the characteristic is a point, provided that it exists. The locus of the characteristic points is called the *edge of regression* of the system (6.4–1).

A characteristic is not always present. Consider for example a set of parallel hyperplanes. A necessary condition for the existence of a characteristic is, evidently, the linear independence of the vectors $\mathrm{d}_q^{\ h} p$, $h = 0, \ldots, m$.

Suppose now that the characteristic points constitute a space curve $x(q)$. This vector function satisfies identically the equations (6.4–6). By differentiating we get

$$\mathrm{d}_q x\, \mathrm{d}_q^{\ h+1} p + x\, \mathrm{d}_q^{\ h+1} p = \mathrm{d}_q^{\ h+1} \eta, \qquad h = 0, \ldots, r-1,$$

whence

$$\mathrm{d}_q x\, \mathrm{d}_q^{\ h} p = 0, \qquad h = 0, \ldots, r-1.$$

Again

$$\mathrm{d}_q^{\ h} x\, \mathrm{d}_q^{\ h} p + \mathrm{d}_q x\, \mathrm{d}_q^{\ h+1} p = 0$$

or

$$\mathrm{d}_q^2 x\, \mathrm{d}_q^{\ h} p = 0, \qquad h = 0, \ldots, r-2.$$

Proceeding along these lines we find

$$\mathrm{d}_q^{\ m} x\, \mathrm{d}_q^{\ h} p = 0, \qquad m = 1, \ldots, r, \qquad h = 0, \ldots, r-m. \qquad (6.4\text{–}7)$$

These equations have an interesting consequence. The m-dimensional osculating space at a point of the curve $x(q)$ can be represented by the vector

$$x + \mathrm{d}_q x \underset{0}{\xi} + \ldots + \mathrm{d}_q^{\ m} x \underset{m-1}{\xi}\,. \qquad (6.4\text{–}8)$$

On account of (6.4–6) and (6.4–7) the product of this vector and either of the vectors $\mathrm{d}_q^{\ h} p$, $h = 0, \ldots, r-m$ vanishes. Hence, *the m-dimensional osculating space of the curve* $x(s)$ *coincides with the m-dimensional characteristic of the system* (6.5–1).

6.4.2 – Developables associated with a curve

In the following example the existence of a curve consisting of characteristic points is granted and also, therefore, that of the locus of the m-dimensional characteristics ($m > 0$), the so-called *developable* of the system (6.4–1). Consider the set of hyperplanes

$$(y - x(s))\underset{r}{u}(s) = 0, \qquad (6.4\text{–}9)$$

the r-dimensional osculating spaces. The $(r-1)$-dimensional characteristic is the intersection of (6.4–9) and

$$(y-x)\mathrm{d}_s \underset{r}{u} = 0$$

or, taking account of Frenet's equations,

$$(y-x) \underset{r-1}{u} = 0, \qquad (6.4\text{–}10)$$

for $\underset{r}{\kappa} \neq 0$. Again the $(r-2)$-dimensional characteristic is the intersection of (6.4–9), (6.4–10) and

$$(y-x)(-\underset{r-2}{u}\,\underset{r-1}{\kappa} + \underset{r}{u}\,\underset{r}{\kappa}) = 0$$

or

$$(y-x) \underset{r-2}{u} = 0. \qquad (6.4\text{–}11)$$

Proceeding in this way we see that the characteristic points satisfy the equations

$$(y-x)\underset{h}{u} = 0, \qquad h = 0, \ldots, r. \qquad (6.4\text{–}12)$$

and the vectors $\underset{0}{u}, \ldots, \underset{r}{u}$ are actually independent. The relations (6.4–12) are only satisfied for $y = x(s)$, that is to say:

The edge of regression of the system of the osculating hyperplanes of a curve is the curve itself.

Another interesting set of hyperplanes connected with a curve $x(s)$ is the system of normal hyperplanes characterized by

$$(y-x)\underset{0}{u} = 0. \qquad (6.4\text{–}13)$$

The $(r-1)$-dimensional characteristic is determined by (6.4–13) and

$$-\underset{0}{u}\,\underset{0}{u} + (y-x)\underset{1}{u}\,\underset{1}{\kappa} = 0$$

or

$$(y-x)\underset{1}{u} = \underset{1}{\rho},$$

where $\underset{1}{\rho}$ stands for $\underset{1}{\kappa}^{-1}$.

In the sequel we shall agree on writing $\underset{h}{\rho}$, rather than $\underset{h}{\kappa}^{-1}$, $h = 1, \ldots, r$. It will be our aim to show that the characteristics of the system (6.4–13) are given by

$$(\mathbf{y}-\mathbf{x})\mathbf{u}_h = \xi_h, \qquad h = 0, 1, \ldots, \mathrm{r}, \tag{6.4–14}$$

where ξ_h denotes a function of $\rho_1, \ldots, \rho_h$, which are in turn functions of the arc length s.

The assertion is true if $h = 0$, provided we take $\xi_0 = 0$. It is also true for $h = 1$, provided we take $\xi_1 = \rho_1$. Let $0 < h < r$ denote a number such that (6.4–14) is true. We find by differentiation

$$-\mathbf{u}_0\mathbf{u}_h + (\mathbf{y}-\mathbf{x})(-\mathbf{u}_{h-1}\kappa_h + \mathbf{u}_{h+1}\kappa_{h+1}) = \mathrm{d}_s\xi_h,$$

and an easy calculation shows that

$$(\mathbf{y}-\mathbf{x})\,\mathbf{u}_{h+1} = \xi_{h+1},$$

with

$$\boxed{\xi_{h+1} = \xi_{h-1}\,\rho_{h+1}/\rho_h + \rho_h\,\mathrm{d}_s\xi_h,} \qquad h = 1, \ldots, \mathrm{r}-1. \tag{6.4–15}$$

These are recursive equations, which determine the functions $\xi_1, \ldots, \xi_\mathrm{r}$ uniquely, since $\xi_1 = \rho_1$.

The edge of regression is the curve

$$\mathbf{y}(s) = \mathbf{x}(s) + \sum_{h=1}^{\mathrm{r}} \mathbf{u}_h\xi_h, \tag{6.4–16}$$

provided the vector on the right does not reduce to a constant.

6.4.3 – The osculating hypersphere of a curve

A problem which is closely related to the considerations included in the second part of the previous section is the following.

Suppose we are given a hypersphere

$$(\mathbf{y}-\mathbf{a})(\mathbf{y}-\mathbf{a}) = \rho^2, \tag{6.4–17}$$

that is, a set of points whose distance from a point $\mathbf{a}$ is equal to ρ. The point $\mathbf{a}$ is the centre of the hypersphere.

A curve $\mathbf{x}(s)$ and the hypersphere (6.5–17) have a point in common if the function

$$f(s) = (\mathbf{x}(s)-\mathbf{a})(\mathbf{x}(s)-\mathbf{a})-\rho^2 \tag{6.4–18}$$

has a zero. We shall say that the hypersphere and the curve have $h+1$ *consecutive points* in common, if the multiplicity of this zero is h. This requires that the zero s also satisfies the equations

$$d_s f = d_s^2 f = \ldots = d_s^h f = 0. \tag{6.4–19}$$

In particular, let h be equal to $r+1$. We then say that the hypersphere *osculates* the curve at the point under consideration. Otherwise stated: it has there $r+2$ consecutive points in common with the curve.

It is easy to show that this hypersphere is uniquely determined. In fact:

$$d_s f = 0$$

implies

$$(\boldsymbol{x}-\boldsymbol{a})\, \underset{0}{\boldsymbol{u}} = 0.$$

The condition

$$d_s^2 f = 0$$

ensures

$$(\boldsymbol{a}-\boldsymbol{x})\, \underset{1}{\boldsymbol{u}} = \underset{1}{\rho}.$$

Proceeding along these lines we easily find that

$$(\boldsymbol{a}-\boldsymbol{x})\underset{h}{\boldsymbol{u}} = \underset{h}{\xi}, \qquad h = 0, \ldots, \mathrm{r}-1, \tag{6.4–19}$$

where the functions $\underset{h}{\xi}$ satisfy the recursive relations (6.4–15). As a consequence we have:

The centre of the osculating hypersphere is given by the coordinate vector

$$\boldsymbol{a} = \boldsymbol{x} + \sum_{h=1}^{\mathrm{r}} \underset{h}{\boldsymbol{u}}\, \underset{h}{\xi}, \tag{6.4–20}$$

where $\underset{1}{\xi}, \ldots, \underset{\mathrm{r}}{\xi}$ *are defined by the recursive relations* (6.4–15) *and* $\underset{1}{\xi} = \rho$. *The radius of this hypersphere is determined by*

$$c^2 = \sum_{h=1}^{\mathrm{r}} \underset{h}{\xi}^2. \tag{6.4–21}$$

We may also state this result in the following form:

The characteristic points of the system of normal hyperplanes of a curve are the centres of the osculating spheres of the curve.

In this connection it is worth while to investigate under which conditions the set of characteristic points reduces to a single point: Then we must have in (6.4–20)

$$d_s \boldsymbol{a} = \boldsymbol{o}.$$

This is, however, the same problem as that dealt with in section 6.2.4, for the equation (6.4–20) is the same as the equation (6.2–31). Thus we find that the function $\underset{1}{\xi}, \ldots, \underset{r}{\xi}$ must satisfy the conditions (6.2–29). As a consequence the radius of the osculating sphere is also constant. This means, however, that the curve is hyperspherical.

It is interesting to observe that the relations (6.2–29) are satisfied by those functions (6.4–15) on which the additional condition

$$0 = \underset{r-1}{\xi} + \underset{r}{\rho}\, \mathrm{d}_s \underset{r}{\xi} \tag{6.4–22}$$

has been imposed.

We conclude this section by discussing a curve in a space whose dimension is an odd number, the curve having the property that the ratios

$$\underset{1}{\kappa}/1,\ \underset{3}{\kappa}/\underset{2}{\kappa},\ \ldots,\ \underset{r}{\kappa}\,/\,\underset{r-1}{\kappa} \tag{6.4–23}$$

are constant.

It follows from (6.4–15) that the coefficients ξ occurring in (6.4–20) are all constants, and indeed those having an even subscript are zero. In fact, first we have

$$\underset{1}{\xi} = \underset{1}{\rho}.$$

On account of (6.4–15) we may infer that

$$\underset{2}{\xi} = 0$$

and

$$\underset{3}{\xi} = \underset{3}{\rho}\, \underset{1}{\xi} / \underset{2}{\rho}$$

is again constant. Assuming that $\underset{h-1}{\xi} = 0$ and $\underset{h}{\xi}$ is constant we find that $\underset{h+1}{\xi} = 0$, while $\underset{h-1}{\xi} =$ constant and $\underset{h}{\xi} = 0$ implies that $\underset{h+1}{\xi} =$ constant. Since the coefficients with odd index are positive the equation (6.4–22) does not hold. Hence the locus of the centre of the osculating hypersphere

$$\overset{*}{x} = x + \underset{1}{u}\underset{1}{\xi} + \underset{3}{u}\underset{3}{\xi} + \ldots + \underset{r-1}{u}\ \underset{r-1}{\xi} \tag{6.4–24}$$

does not reduce to a point. In addition we find that *the radius of the osculating hypersphere is constant along the curve.*

Differentiating (6.4–24) we get

$$\begin{aligned}\mathrm{d}_s \overset{*}{\boldsymbol{x}} &= \boldsymbol{u}_0 + (-\boldsymbol{u}_0 \kappa_1 + \boldsymbol{u}_2 \kappa_2)\xi_2 + (-\boldsymbol{u}_2 \kappa_3 + \boldsymbol{u}_4 \kappa_4)\xi_4 + \ldots + (-\boldsymbol{u}_{r-2}\kappa_{r-1} + \boldsymbol{u}_r \kappa_r)\xi_{r-1} \\ &= \boldsymbol{u}_0(1-\xi_1\kappa_1) + \boldsymbol{u}_2(\xi_1\kappa_2 - \xi_3\kappa_3) + \ldots + \boldsymbol{u}_{r-2}(\xi_{r-3}\kappa_{r-2} - \xi_{r-1}\kappa_{r-1}) + \boldsymbol{u}_r\xi_{r-1}\kappa_r.\end{aligned}$$

But from (6.4–15) we deduce

$$\xi_{h+1}\kappa_{h+1} - \xi_{h-1}\kappa_h = 0, \quad h = 1, \ldots, r-1. \tag{6.4–25}$$

Hence, since $\xi_0 = 0$,

$$\mathrm{d}_s \overset{*}{\boldsymbol{x}} = \boldsymbol{u}_r \xi_{r-1} \kappa_r. \tag{6.4–26}$$

Denoting by $\overset{*}{s}$ the arc length of the curve (6.4–24) we evidently have

$$(\mathrm{d}_s \overset{*}{s})^2 = \xi_{r-1}^2 \kappa_r^2.$$

We agree to take

$$\mathrm{d}_s \overset{*}{s} = -\xi_{r-1}\kappa_r. \tag{6.4–27}$$

Then it follows that

$$\overset{*}{\boldsymbol{u}}_0 = \mathrm{d}_{\overset{*}{s}} \overset{*}{\boldsymbol{x}} = \mathrm{d}_s \overset{*}{\boldsymbol{x}} \frac{1}{\mathrm{d}_s \overset{*}{s}} = -\boldsymbol{u}_r \xi_{r-1}\kappa_r \,/\, \xi_{r-1}\kappa_r = -\boldsymbol{u}_r. \tag{6.4–28}$$

Next we have

$$\mathrm{d}_{\overset{*}{s}} \overset{*}{\boldsymbol{u}}_0 = \overset{*}{\boldsymbol{u}}_1 \overset{*}{\kappa}_1 = \mathrm{d}_s \boldsymbol{u}_r / \xi_{r-1}\kappa_r = -\boldsymbol{u}_{r-1}\kappa_r / \xi_{r-1}\kappa_r = -\boldsymbol{u}_{r-1} \frac{1}{\xi_{r-1}}.$$

Observing that $\overset{*}{\boldsymbol{u}}_1$ and $\boldsymbol{u}_{r-1}$ are unit vectors and that ξ_{r-1} is positive we may infer that

$$\overset{*}{\boldsymbol{u}}_1 = -\boldsymbol{u}_{r-1}.$$

Thus we arrive at the conjecture that in general

$$\overset{*}{\boldsymbol{u}}_h = -\boldsymbol{u}_{r-h}, \qquad h = 0, \ldots, r. \tag{6.4–29}$$

Assume that this conjecture has already been verified for all subscripts from 0 up to a certain index h. By differentiating we get

$$d_{\overset{*}{s}} \overset{*}{\underset{h}{u}} = - \overset{*}{\underset{h-1}{u}} \overset{*}{\underset{h}{\kappa}} + \overset{*}{\underset{h+1}{u}} \overset{*}{\underset{h+1}{\kappa}} = (\underset{r-h-1}{u} \underset{r-h}{\kappa} - \underset{r-h+1}{u} \underset{r-h+1}{\kappa}) \frac{1}{\underset{r-1}{\xi} \underset{r}{\kappa}}. \tag{6.4–30}$$

Observing that $\overset{*}{\underset{h-1}{u}}, \overset{*}{\underset{h+1}{u}}$ are independent unit vectors as are $\underset{r-h-1}{u}, \underset{r-h+1}{u}$ and that by hypothesis $\overset{*}{\underset{h-1}{u}} = - \underset{r-h+1}{u}$, we find that also

$$\overset{*}{\underset{h+1}{u}} = - \underset{r-h-1}{u}.$$

This proves the assertion.

Next we introduce the numbers

$$\overset{*}{\underset{h}{\xi}} = \underset{r-h}{\xi}, \qquad h = 1, 3, \ldots, r-1.$$

Then (6.4–24) may be written in the form

$$x = \overset{*}{x} + \overset{*}{\underset{1}{u}} \underset{1}{\xi} + \overset{*}{\underset{3}{u}} \underset{3}{\xi} + \ldots + \overset{*}{\underset{r-1}{u}} \underset{r-1}{\xi}. \tag{6.4–31}$$

From (6.4–30) we also infer that

$$\overset{*}{\underset{h}{\kappa}} = \underset{r-h}{\kappa} / \underset{r-1}{\xi} \underset{r}{\kappa},$$

i.e.,

$$\overset{*}{\underset{h}{\kappa}} / \overset{*}{\underset{h+1}{\kappa}} = \underset{r-h}{\kappa} / \underset{r-h-1}{\kappa}.$$

This means that the ratios

$$\overset{*}{\underset{1}{\kappa}}/1, \quad \overset{*}{\underset{3}{\kappa}}/\overset{*}{\underset{2}{\kappa}}, \ldots, \overset{*}{\underset{r}{\kappa}}/\overset{*}{\underset{r-1}{\kappa}}.$$

are constant and that the numbers $\overset{*}{\underset{h}{\xi}}$ satisfy equations similar to (6.4–15) satisfied by the $\underset{h}{\xi}$. As a consequence the curve (6.4–31) is the locus of the centre of the osculating hypersphere of the curve (6.4–24). Thus we proved:

If for a curve $x(s)$ *the ratios* (6.2–24) *are constant, where* r *is an even number, then a similar condition holds for the locus* $\overset{*}{x}(s)$ *of the centre of the osculating sphere, and the locus of the centre of the osculating hypersphere of the curve* $\overset{*}{x}$ *is just the original curve* x.

6.5 – Intrinsic equations

6.5.1 – The existence of a curve with prescribed curvatures

In section 6.2.1 we found that a curve in an $(r+1)$-dimensional space (and not in a space of lower dimension) possesses r curvatures $\underset{1}{\kappa}(s), \ldots, \underset{r}{\kappa}(s)$ which are positive functions (except for isolated values of s). In this section we shall establish the theorem stating that these curvature functions are not subject to any further restriction and that they determine the curve uniquely, if we do not wish to distinguish between two curves of which one is obtained from the other by a simple displacement. More precisely stated:

In an $(r+1)$-dimensional space there exists one and only one curve such that

(i) *the curve passes through a given point;*

(ii) *at this point the frame consisting of the tangent vector and the normal vectors coincides with a given orthonormal frame;*

(iii) *the curvatures are equal to given continuous non-vanishing and non-negative functions*:

$$\underset{1}{\kappa} = \underset{1}{k}(s), \ldots, \underset{r}{\kappa} = \underset{r}{k}(s); \tag{6.5–1}$$

(iv) *the variable s is the arc length of the curve.*

The equations relating the arc length and the curvatures of the curve are called the *intrinsic equations* of the curve.

In order to prove the theorem we consider the following system of ordinary differential equations for the functions $\underset{0}{u}, \ldots, \underset{r}{u}$:

$$\begin{aligned} d_s \underset{0}{u} &= \underset{1}{u}\,\underset{1}{k}, \\ d_s \underset{h}{u} &= -\underset{h-1}{u}\,\underset{h}{k} + \underset{h+1}{u}\,\underset{h+1}{k}, \qquad h = 1, \ldots, r-1, \\ d_s \underset{r}{u} &= -\underset{r-1}{u}\,\underset{r}{k}, \end{aligned} \tag{6.5–2}$$

which present the same pattern as Frenet's equations.

From the existence theorem we may conclude that this system possesses a uniquely determined set of solutions taking prescribed values at $s = a$, the number a being a given constant.

Next we select $r+1$ systems of solutions

$$\underset{k0}{u}, \underset{k1}{u}, \ldots, \underset{kr}{u}, \qquad k = 0, \ldots, r, \tag{6.5–3}$$

where $\underset{mn}{u}$ takes the initial value 0 if $m \neq n$, and 1 if $m = n$. Additionally we introduce functions v_{mn} defined by

$$v_{mn} = \sum_{k=0}^{r} u_{kn} u_{km}, \qquad m, n = 0, \ldots, r. \tag{6.5–4}$$

An easy calculation shows that these functions satisfy the system of differential equations

$$d_s v_{mn} = - v_{m-1\,n} k_m + v_{k+1\,m} k_{m+1} - v_{m\,n-1} k_n + v_{m\,n+1} k_{n+1}, \tag{6.5–5}$$

where it is understood that v_{mn} is identically zero whenever $m = -1$ or $n = -1$, while $k_{r+1} = 0$.

A solution of this system is given by $v_{mn} = 0$ if $m \neq n$ and $v_{mn} = 1$ if $m = n$. These values are, however, the initial values of the functions (6.5–4). Hence, on account of the assertion of uniqueness in the existence theorem the functions (6.5–4) coincide with those evaluated above.

Now let

$$\overset{*}{u}_0, \ldots, \overset{*}{u}_r \tag{6.5–6}$$

denote a given orthonormal frame. We introduce the vector functions

$$u_h = \sum_{k=0}^{r} \overset{*}{u}_k u_{kh}, \qquad h = 0, \ldots, r. \tag{6.5–7}$$

Then, evidently,

$$u_m u_n = \sum_{k=0}^{r} u_{km} u_{kn} = v_{mn}, \qquad m, n = 0, \ldots, r, \tag{6.5–8}$$

and this means that the vectors (6.5–7) constitute an orthonormal frame along the curve coinciding with the frame (6.5–6) for $s = a$.

Now we are sufficiently prepared to prove the assertion that the desired curve is represented by

$$x(s) = \int_a^s u_0 \, ds + x(a), \tag{6.5–9}$$

the integral denoting a vector function whose derivative is u_0, and which vanishes for $s = a$. Moreover $x(a)$ denotes the coordinate vector of the given point mentioned in part (i) of the theorem. In fact, first we have:

$$d_s x = u_0.$$

Hence

$$\mathrm{d}_s\boldsymbol{x}\,\mathrm{d}_s\boldsymbol{x} = \underset{0}{\boldsymbol{u}}\,\underset{0}{\boldsymbol{u}} = \underset{00}{v} = 1$$

and it follows that s is the arc length of the curve (6.5–9). Proceeding, we find in view of (6.5–7) and (6.5–2)

$$\mathrm{d}_s\underset{h}{\boldsymbol{u}} = \sum_{k=0}^{r} \overset{*}{\underset{k}{\boldsymbol{u}}}\,\mathrm{d}_s\underset{kh}{u} = \sum_{k=0}^{r} \overset{*}{\underset{k}{\boldsymbol{u}}}\,(-\underset{k\,h-1}{u}\,\underset{h}{k} + \underset{k\,h+1}{u}\,\underset{h+1}{k})$$

$$= -\underset{h-1}{\boldsymbol{u}}\,\underset{h}{k} + \underset{h+1}{\boldsymbol{u}}\,\underset{h+1}{k}, \qquad h = 1, \ldots, r-1,$$

completed with

$$\mathrm{d}_s\underset{0}{\boldsymbol{u}} = \underset{1}{\boldsymbol{u}}\,\underset{1}{k}, \qquad \mathrm{d}_s\underset{r}{\boldsymbol{u}} = -\underset{r-1}{\boldsymbol{u}}\,\underset{r}{k}.$$

Hence the functions (6.5–1) are precisely the curvatures of the curve (6.5–9) and this completes the proof of the theorem.

6.5.2 – A THEOREM ABOUT INVOLUTES

In section 6.3.1 we proved that two involutes cut a segment of constant length on each of the generators of the tangent surface of the given curve. This result may be generalized to the case of involutes of higher order in the following fashion:

The configuration of the traces of the involutes of order n on the n-dimensional osculating spaces of a curve $\boldsymbol{x}(s)$ *moves as a rigid body as s varies.*

In an auxiliary n-dimensional flat we take a curve $\boldsymbol{z}(s)$ having its curvatures equal to the first $n-1$ curvatures $\underset{1}{\kappa}, \ldots, \underset{n-1}{\kappa}$ of the given curve $\boldsymbol{x}(s)$, s being also the arc length of this new curve. This is possible on account of the theorem of the previous section. Let $\underset{0}{\boldsymbol{v}}$ denote the tangent vector and $\underset{1}{\boldsymbol{v}}, \ldots, \underset{n-1}{\boldsymbol{v}}$ its normals. Next we consider the vector function

$$\overset{*}{\boldsymbol{z}}(s) = \boldsymbol{z}(s) + \sum_{h=0}^{n-1} \underset{h}{\boldsymbol{v}}\,\underset{h}{\xi}, \tag{6.5–10}$$

where $\underset{0}{\xi}, \ldots, \underset{n-1}{\xi}$ are solutions of the system of differential equations (6.3–6).

On differentiating we see at once that

$$\mathrm{d}_s\overset{*}{\boldsymbol{z}} = \boldsymbol{o} \tag{6.5–11}$$

identically. Hence $\overset{*}{\boldsymbol{z}}$ is a constant vector and determines an invariable point. This is, however, the content of the assertion.

6.5.3 – The applicability of a developable

Let $\boldsymbol{x}(s)$ denote a given curve in an $(r+1)$-dimensional space. The n-dimensional characteristics of the system of the osculating hyperplanes of this curve are represented by

$$\boldsymbol{y} = \boldsymbol{x} + \underset{0}{\boldsymbol{u}}\,\underset{0}{\xi} + \ldots + \underset{n-1}{\boldsymbol{u}}\;\underset{n-1}{\xi}\,. \tag{6.5-12}$$

They generate an $(n+1)$-dimensional manifold referred to parameters

$$q^1 = s, \qquad q^2 = \underset{0}{\xi}, \ldots, q^{n+1} = \underset{n-1}{\xi}\,. \tag{6.5-13}$$

The tangent vectors of this manifold are respectively

$$\partial_1 \boldsymbol{y} = \underset{0}{\boldsymbol{u}} + \underset{1}{\boldsymbol{u}}\,\underset{0}{\xi}\,\underset{1}{\kappa} + (-\underset{0}{\boldsymbol{u}}\,\underset{1}{\kappa} + \underset{2}{\boldsymbol{u}}\,\underset{2}{\kappa})\,\underset{1}{\xi} + \ldots + (-\underset{n-2}{\boldsymbol{u}}\;\underset{n-1}{\kappa} + \underset{n}{\boldsymbol{u}}\,\underset{n}{\kappa})\;\underset{n-1}{\xi}\,,$$

$$\partial_2 \boldsymbol{y} = \underset{0}{\boldsymbol{u}}, \qquad \partial_3 \boldsymbol{y} = \underset{1}{\boldsymbol{u}}, \ldots, \qquad \partial_{n+1} \boldsymbol{y} = \underset{n-1}{\boldsymbol{u}}\,.$$

As a consequence the components of the metric tensor involve only the curvatures $\underset{1}{\kappa}, \ldots, \underset{n}{\kappa}$ and are uniquely determined by these curvatures.

Now we refer to the theorem of section 6.5.1. It states that we can find in an $(n+1)$-dimensional space a curve having the same curvatures as the given curve $\boldsymbol{x}(s)$. But the n-dimensional osculating spaces sweep through an $(n+1)$-dimensional part of this space, being an $(n+1)$-dimensional flat referred to the same parameters (6.5–13) as is the developable considered above. Since the metric tensor of this flat is exactly the same as that of the developable (this flat being considered as the developable of the auxiliary curve in the $(n+1)$-dimensional space) we may infer that:

An $(n+1)$-dimensional developable associated with a curve is isometrically equivalent to an $(n+1)$-dimensional flat.

This fact explains the name "developable".

CHAPTER 7

GEODESIC DIFFERENTIATION

By a one-dimensional vector field on a manifold we mean a system of tangential vectors, depending on a parameter q. The derivative of this field measures the rate of change with respect to q. But in most cases the derivative is a vector which is not tangential and, therefore, cannot be characterized by components with respect to a basis that spans a tangent space. We extricate ourselves from this difficulty by considering the projection of the derivative on the tangent space. Thus we arrive at a modified process of differentiation, which turns out to be very useful. It has, moreover, an intrinsic character, that is to say, when we consider two corresponding fields on manifolds in isometric correspondence, then the modified derivatives also correspond.

This process of differentiation serves among other things to define a type of curves, which generalize the notion of straight lines in Euclidean space. These curves will be referred to as geodesic lines or paths. Thus a considerable part of this chapter will be devoted to the study of these curves.

7.1 – The geodesic derivative

7.1.1 – Definition

On a manifold $\boldsymbol{x}(q^\kappa)$, $\kappa = 1, \ldots, n$, $n \geqq 2$, we consider a curve characterized by a set of functions

$$q^\kappa(q). \tag{7.1–1}$$

As usual we write briefly $\boldsymbol{x}(q)$ instead of $\boldsymbol{x}(q^1(q), \ldots, q^n(q))$.

Next we consider a field of tangential vectors

$$\boldsymbol{v}(q) \tag{7.1–2}$$

defined along the curve. That means: to every point q of the curve corresponds a vector $\boldsymbol{v}(q)$ included in the tangent space of the manifold at this point. We are only concerned with fields which are continuously differentiable with respect to the variable q. In general, however, the derivative $\mathrm{d}_q \boldsymbol{v}$ is a vector which is not included in the tangent space at the point under consideration. The projection of this vector will present itself as a very important entity. It will be referred to as the *geodesic derivative* of the field $\boldsymbol{v}(q)$. We shall denote it by $\mathsf{D}_q \boldsymbol{v}(q)$, i.e.,

$$\mathsf{D}_q \boldsymbol{v}(q) = \mathsf{P}\mathrm{d}_q \boldsymbol{v}(q), \tag{7.1–3}$$

where P denotes the projection operator associated with the tangent space at the point under consideration (see section 2.2.1).

7.1.2 – Elementary rules

It is an easy matter to show that many rules of elementary calculus are also valid for the geodesic derivative.

Let $\alpha(q)$ denote a function of the variable q, i.e., a function defined along the given curve. Since P is a linear operator, we deduce from

$$\mathrm{d}_q(\boldsymbol{v}\alpha) = \mathrm{d}_q \boldsymbol{v}\alpha + \boldsymbol{v}\mathrm{d}_q \alpha$$

that

$$\mathsf{P}\mathrm{d}_q(\boldsymbol{v}\alpha) = \mathsf{P}\mathrm{d}_q \boldsymbol{v}\alpha + (\mathsf{P}\boldsymbol{v})\mathrm{d}_q \alpha$$

or

$$\boxed{\mathsf{D}_q(\boldsymbol{v}\alpha) = \mathsf{D}_q \boldsymbol{v}\,\alpha + \boldsymbol{v}\mathrm{d}_q \alpha,} \tag{7.1–4}$$

for $\mathsf{P}\boldsymbol{v} = \boldsymbol{v}$.

Next we consider two tangential fields $\boldsymbol{v}(q)$ and $\boldsymbol{w}(q)$ defined along the given curve. From

$$\mathrm{d}_q(\boldsymbol{v}+\boldsymbol{w}) = \mathrm{d}_q \boldsymbol{v} + \mathrm{d}_q \boldsymbol{w}$$

we deduce

$$\mathsf{P}\mathrm{d}_q(\boldsymbol{v}+\boldsymbol{w}) = \mathsf{P}\mathrm{d}_q \boldsymbol{v} + \mathrm{d}_q \boldsymbol{w}$$

i.e.,

$$\boxed{\mathsf{D}_q(\boldsymbol{v}+\boldsymbol{w}) = \mathsf{D}_q \boldsymbol{v} + \mathsf{D}_q \boldsymbol{w}.} \tag{7.1–5}$$

Differentiating the scalar product $\boldsymbol{vw}$ we get

$$\mathrm{d}_q(\boldsymbol{vw}) = (\mathrm{d}_q \boldsymbol{v})\boldsymbol{w} + \boldsymbol{v}\mathrm{d}_q \boldsymbol{w}.$$

Since $\mathrm{d}_q \boldsymbol{v} - \mathsf{P}\mathrm{d}_q \boldsymbol{v}$ is orthogonal to the tangent space and $\boldsymbol{w}$ is included in this space, we evidently have $(\mathrm{d}_q \boldsymbol{v} - \mathsf{P}\mathrm{d}_q \boldsymbol{v})\boldsymbol{w} = 0$, i.e., $(\mathrm{d}_q \boldsymbol{v})\boldsymbol{w} = (\mathsf{P}\mathrm{d}_q \boldsymbol{v})\boldsymbol{w} = (\mathsf{D}_q \boldsymbol{v})\boldsymbol{w}$. By the same arguments we conclude that also $\boldsymbol{v}\mathrm{d}_q \boldsymbol{w} = \boldsymbol{v}\mathsf{D}_q \boldsymbol{w}$. Hence

$$\boxed{\mathrm{d}_q(\boldsymbol{vw}) = (\mathsf{D}_q \boldsymbol{v})\boldsymbol{w} + \boldsymbol{v}\mathsf{D}_q \boldsymbol{w}.} \tag{7.1–6}$$

It should be noticed that $\mathsf{D}_q \boldsymbol{v} = \boldsymbol{o}$ along the curve means that $\mathrm{d}_q \boldsymbol{v}$ is everywhere orthogonal to the manifold. Hence $\mathsf{D}_q \boldsymbol{v} = \boldsymbol{o}$ does not imply $\boldsymbol{v} = \boldsymbol{o}$.

Finally we wish to check that the ordinary chain rule holds when we take another parameter along the curve. In fact, if $q(p)$ is a function of p then

$$\mathrm{d}_p \boldsymbol{v} = \mathrm{d}_q \boldsymbol{v}\, \mathrm{d}_p q$$

whence, on applying the operator P,

$$\boxed{\mathsf{D}_p \boldsymbol{v} = \mathsf{D}_q \boldsymbol{v}\, \mathrm{d}_p q.} \tag{7.1–7}$$

7.1.3 – Frenet's Equations

Proceeding as in section 6.2.1 we can get a system of equations which are to be considered as a generalization of Frenet's equations (6.2–7) for curves on general manifolds. We may identify $\mathsf{D}_s \boldsymbol{x}$ with $\mathrm{d}_s \boldsymbol{x}$, for it is a tangential vector. On applying the operator D_s repeatedly we obtain a set of vectors

$$\mathsf{D}_s \boldsymbol{x},\ \mathsf{D}_s^2 \boldsymbol{x},\ \ldots,\ \mathsf{D}_s^n \boldsymbol{x}, \tag{7.1–8}$$

all belonging to the tangent space of the manifold at the point where (7.1–8) are evaluated. We consider only the case that these vectors are linearly independent. On applying the Gram-Schmidt process we can find a system of mutually orthogonal unit vectors

$$\underset{0}{\boldsymbol{u}},\ \underset{1}{\boldsymbol{u}},\ \ldots,\ \underset{n-1}{\boldsymbol{u}} \tag{7.1–9}$$

which also span the tangent space. In just the same way as in section 6.2.1 we find Frenet's equations

$$\boxed{\begin{aligned} \mathsf{D}_s \underset{0}{\boldsymbol{u}} &= \underset{1}{\boldsymbol{u}}\,\underset{1}{\tilde{\kappa}}, \\ \mathsf{D}_s \underset{h}{\boldsymbol{u}} &= -\underset{h-1}{\boldsymbol{u}}\,\underset{h}{\tilde{\kappa}} + \underset{h+1}{\boldsymbol{u}}\,\underset{h+1}{\tilde{\kappa}}, \\ \mathsf{D}_s \underset{n-1}{\boldsymbol{u}} &= -\underset{n-2}{\boldsymbol{u}}\,\underset{n-1}{\tilde{\kappa}}, \end{aligned}} \quad h = 1, \ldots, n-2. \tag{7.1–10}$$

The functions $\underset{1}{\tilde{\kappa}}, \underset{2}{\tilde{\kappa}}, \ldots, \underset{n-1}{\tilde{\kappa}}$ are called the *geodesic curvature* and the *geodesic torsions* just as in the case of a curve in linear space. They may be evaluated by formulas similar to those derived in section 6.2.2.

The geodesic curvature may be defined independent of the assumptions underlying the derivation of Frenet's equations. In fact, it follows from the first equation of (7.1–10) that $\underset{1}{\tilde{\kappa}}$ *is the absolute value of the vector* $\mathsf{D}_s \underset{0}{\boldsymbol{u}} = \mathsf{D}_s \mathrm{d}_s \boldsymbol{x}$. Then the case that it vanishes is not excluded.

7.1.4 – Geodesically parallel vectors

A line segment in ordinary space undergoes a parallel displacement when it moves in such a way that its direction remains unaltered. Hence the rate of change in every position can be represented by a vector proportional to the vector defined by the segment in that position.

This idea of parallelism can easily be extended to a vector field on a general manifold, provided that we define the rate of change by means of the geodesic derivative.

We shall say that a non-zero vector field $\boldsymbol{v}(q)$ defined along a curve $q^\kappa(q)$, $\kappa = 1, \ldots, n$ on the manifold $\boldsymbol{x}(q^\kappa)$ consists of *geodesically parallel vectors,* if we can find a function $\alpha(q)$ such that

$$\mathsf{D}_q \boldsymbol{v}(q) = \boldsymbol{v}(q)\alpha(q) \tag{7.1–11}$$

along the curve. It follows from (6.1–7) that this notion is independent of the parameter of the given curve.

A field $\boldsymbol{v}(q)$ *of geodesically parallel vectors along a curve remains a field of the same kind, when the vector function* $\boldsymbol{v}(q)$ *is multiplied by a function* $\lambda(q)$ *which vanishes nowhere.*

This follows at once from (7.1–4) and (7.1–11), for

$$\begin{aligned}\mathsf{D}_q(\boldsymbol{v}\lambda) &= (\mathsf{D}_q \boldsymbol{v})\lambda + \boldsymbol{v}\mathrm{d}_q\lambda = \boldsymbol{v}\alpha\lambda + \boldsymbol{v}\mathrm{d}_q\lambda \\ &= \boldsymbol{v}\lambda(\alpha + \mathrm{d}_q\lambda/\lambda).\end{aligned}$$

The next theorem expresses a remarkable analogy between the parallelism on a manifold and the parallelism in ordinary space.

If $\boldsymbol{v}(q)$ *and* $\boldsymbol{w}(q)$ *are vector fields, each consisting of geodesically parallel vectors different from zero along a given curve, then the angle between* $\boldsymbol{v}(q)$ *and* $\boldsymbol{w}(q)$ *is constant along the curve.*

In view of the previous theorem we may suppose that $\boldsymbol{v}(q)$ and $\boldsymbol{w}(q)$ are vectors of norm unity along the curve. The angle θ between these vectors is determined by

$$\cos\theta = \boldsymbol{v}\boldsymbol{w}. \tag{7.1–12}$$

It follows from (7.1–6) that

$$\mathrm{d}_q \cos\theta = (\mathsf{D}_q \boldsymbol{v})\boldsymbol{w} + \boldsymbol{v}\mathsf{D}_q \boldsymbol{w}. \tag{7.1–13}$$

By hypothesis there exist functions $\alpha(q)$ and $\beta(q)$ such that

$$\mathsf{D}_q \boldsymbol{v} = \boldsymbol{v}\alpha, \qquad \mathsf{D}_q \boldsymbol{w} = \boldsymbol{w}\beta. \tag{7.1–14}$$

We made the assumption that $\boldsymbol{vv} = 1$, $\boldsymbol{ww} = 1$. Hence $\boldsymbol{v}\mathsf{D}_q \boldsymbol{v} = 0$, $\boldsymbol{w}\mathsf{D}_q \boldsymbol{w} = 0$, and on multiplying the members of the equations (7.1–14) by $\boldsymbol{v}$ and $\boldsymbol{w}$ respectively we find that $\alpha = \beta = 0$ identically along the curve. Thus we may infer from (7.1–13) that $\mathrm{d}_q \cos\theta = 0$, i.e., θ is constant along the curve.

7.2 – Geodesic curves

7.2.1 – Definition

It may happen, that the field of tangent vectors of a curve consists of parallel vectors. This fact expresses a particular property of the curve, viz. that the curve is "as straight as possible", for the direction of its tangents does not change along the curve, provided that the rate of change is measured by the geodesic derivative. Thus we may consider a type of curves characterized by the following definition:

We shall say that a curve $\boldsymbol{x}(q)$ given by the functions $q^\kappa(q)$ on a manifold $\boldsymbol{x}(q^\kappa)$ is a *geodesic curve* or a *path*, when its field of tangent vectors is a field of geodesically parallel vectors. This means that there is a function $\alpha(q)$ such that

$$\mathsf{D}_q \mathrm{d}_q \boldsymbol{x}(q) = \mathrm{d}_q \boldsymbol{x}(q)\alpha(q). \tag{7.2–1}$$

With reference to the first theorem of section 7.1.4 we may state that the notion of geodesic curve is independent of the choice of the parameter along the curve.

7.2.2 – Natural parameters

We wish to show that we can simplify the equation (7.2–1) considerably by taking an appropriate parameter along the curve. Let $t(q)$ denote a function whose derivative $\mathrm{d}_q t$ vanishes nowhere. Then we may take t as a parameter and we have

$$\mathrm{d}_q \boldsymbol{x} = \mathrm{d}_t \boldsymbol{x}\, \mathrm{d}_q t,$$
$$\mathsf{D}_q \mathrm{d}_q \boldsymbol{x} = \mathsf{D}_t \mathrm{d}_t \boldsymbol{x} (\mathrm{d}_q t)^2 + \mathrm{d}_t \boldsymbol{x}\, \mathrm{d}_q{}^2 t.$$

The equation (7.2–1) appears in the form

$$\mathsf{D}_t \mathrm{d}_t \boldsymbol{x} (\mathrm{d}_q t)^2 + \mathrm{d}_t \boldsymbol{x} \mathrm{d}_q{}^2 t = \mathrm{d}_t \boldsymbol{x}\, \alpha \mathrm{d}_q t. \tag{7.2–2}$$

Next we impose on t the condition that

$$\mathrm{d}_q{}^2 t = \alpha \mathrm{d}_q t. \tag{7.2–3}$$

Then (7.2–1) takes the simple form

$$\boxed{\mathsf{D}_t \mathrm{d}_t \boldsymbol{x} = \boldsymbol{o},} \tag{7.2–4}$$

where t is any solution of the differential equation (7.2–3). It is easy to see that all solutions of this equation may be written in the form

$$t = as+b, \tag{7.2–5}$$

where $s = s(q)$ is a particular solution of the equation different from zero and a, b are constants. Conversely, when t is a parameter such that the curve characterized by (7.2–1) is also characterized by (7.2–4), then t is a solution of the differential equation (7.2–3). This follows at once from (7.2–2). A parameter of this kind is called a *natural parameter* of the geodesic curve.

The arc length of a geodesic curve is a natural parameter.

In view of the preceding considerations this is almost trivial. In fact, when s denotes the arc length, then

$$\mathrm{d}_s \boldsymbol{x}\, \mathrm{d}_s \boldsymbol{x} = 1 \tag{7.2–6}$$

and consequently

$$\mathsf{D}_s \mathrm{d}_s \boldsymbol{x} \cdot \mathrm{d}_s \boldsymbol{x} = 0. \tag{7.2–7}$$

By hypothesis there is a function $\alpha(s)$ such that

$$\mathsf{D}_s \mathrm{d}_s \boldsymbol{x} = \mathrm{d}_s \boldsymbol{x}\, \alpha, \tag{7.2–8}$$

and on multiplying both members of this equation by $\mathrm{d}_s \boldsymbol{x}$, we find that $\alpha(s) = 0$ along the curve. Hence (7.2–4) holds for $t = s$.

This result can be interpreted in the following way. It was pointed out at the end of section 7.1.3 that the absolute value of the vector $\mathsf{D}_s \mathrm{d}_s \boldsymbol{x}$ is the geodesic curvature of the curve. Hence:

Along a geodesic curve the geodesic curvature is identically zero.

A more geometric interpretation of the result obtained above is the following:

A curve on a manifold is a geodesic curve, if and only if at every point the principal normal of the curve is orthogonal to the manifold.

The principal normal is the direction of the curvature vector $\mathrm{d}_s^2 \boldsymbol{x}$ (see section 6.1.2) provided that it is not the zero vector. As a consequence $\mathsf{D}_s \mathrm{d}_s \boldsymbol{x} = \mathsf{P}\mathrm{d}_s^2 \boldsymbol{x}$ is zero only when $\mathrm{d}_s^2 \boldsymbol{x}$ is orthogonal to the manifold. This argument remains valid when $\mathrm{d}_s^2 \boldsymbol{x}$ is zero in isolated points, provided the normal vector can be defined by continuity. If, however, $\mathrm{d}_s^2 \boldsymbol{x} = \boldsymbol{o}$ identically, then the curve is a straight line and every vector orthogonal to a direction vector may be considered as a principal normal.

7.3 – The Christoffel three index symbols

7.3.1 – Definition

Since at every point of the manifold the vectors

$$\boldsymbol{x}_\kappa \doteq \partial_\kappa \boldsymbol{x}\,(q^1, \ldots, q^n), \qquad \kappa = 1, \ldots, n, \tag{7.3–1}$$

are defined, we always can form a field

$$\boldsymbol{x}_\kappa(q) = \partial_\kappa \boldsymbol{x}\,(q^1\,(q), \ldots, q^n(q)) \tag{7.3–2}$$

along a curve $q^\kappa(q)$, $\kappa = 1, \ldots, n$. Next we apply the process of geodesic differentiation and we evidently have

$$\mathsf{D}_q \boldsymbol{x}_\mu\,(q) = \mathsf{P}\partial_{\lambda\mu}\boldsymbol{x}\,\mathrm{d}_q q^\lambda, \qquad \mu = 1, \ldots, n, \tag{7.3–3}$$

$\partial_{\lambda\mu}\boldsymbol{x}$ denoting the second partial derivative $\partial^2\boldsymbol{x}/\partial q^\lambda\,\partial q^\mu$. In the symbol $\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x}$ are collected n^2 tangential vectors. We may refer them to a frame by writing

$$\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x} = \boldsymbol{x}^\kappa\,\Gamma_{\kappa\lambda\mu} = \boldsymbol{x}_\kappa\,\Gamma^\kappa{}_{\lambda\mu}, \tag{7.3–4}$$

where $\Gamma_{\kappa\lambda\mu}$ are the components of the vector $\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x}$ with respect to the basis $\boldsymbol{x}^\kappa$, and $\Gamma^\kappa{}_{\lambda\mu}$ are the components with respect to the basis $\boldsymbol{x}_\kappa$. These expressions are called the *Christoffel three index symbols* of the first and second kind, respectively.

Since

$$\boldsymbol{x}_\kappa\,\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x} = \boldsymbol{x}_\kappa\,\partial_{\lambda\mu}\boldsymbol{x}, \qquad \boldsymbol{x}^\kappa\,\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x} = \boldsymbol{x}^\kappa\,\partial_{\lambda\mu}\boldsymbol{x},$$

we readily find from (7.3–4)

$$\boxed{\Gamma_{\kappa\lambda\mu} = \boldsymbol{x}_\kappa\,\partial_{\lambda\mu}\boldsymbol{x}} \tag{7.3–5}$$

and

$$\boxed{\Gamma^\kappa{}_{\lambda\mu} = \boldsymbol{x}^\kappa\,\partial_{\lambda\mu}\boldsymbol{x}.} \tag{7.3–6}$$

A direct consequence is expressed by the relations

$$\boxed{\Gamma_{\kappa\lambda\mu} = \Gamma^\rho{}_{\lambda\mu}\,g_{\kappa\rho}, \qquad \Gamma^\kappa{}_{\lambda\mu} = \Gamma_{\rho\lambda\mu}\,g^{\kappa\rho}.} \tag{7.3–7}$$

It is clear that the Christoffel three index symbols are symmetric with respect to the indices λ and μ.

The importance of these symbols lies in the fact that they are isometric invariants, for in the next section we shall show that they can be expressed in terms of the components of the metric tensor and its derivatives.

7.3.2 – Explicit expressions for the Christoffel symbols

On differentiating

$$g_{\lambda\mu} = \partial_\lambda \boldsymbol{x}\, \partial_\mu \boldsymbol{x} \tag{7.3–8}$$

with respect to q^κ, we obtain

$$\partial_\kappa g_{\lambda\mu} = \partial_\lambda \boldsymbol{x}\, \partial_{\mu\kappa} \boldsymbol{x} + \partial_\mu \boldsymbol{x}\, \partial_{\kappa\lambda} \boldsymbol{x},$$

whence, on account of (7.3–5),

$$\partial_\kappa g_{\lambda\mu} = \Gamma_{\lambda\mu\kappa} + \Gamma_{\mu\lambda\kappa}. \tag{7.3–9}$$

On performing a cyclic permutation of the indices we get

$$\partial_\lambda g_{\mu\kappa} = \Gamma_{\mu\kappa\lambda} + \Gamma_{\kappa\mu\lambda}. \tag{7.3–10}$$

Again

$$\partial_\mu g_{\kappa\lambda} = \Gamma_{\kappa\lambda\mu} + \Gamma_{\lambda\kappa\mu}. \tag{7.3–11}$$

By adding corresponding members of (7.3–10) and (7.3–11) and subtracting from the result obtained corresponding members of (7.3–9) we readily find

$$\boxed{\Gamma_{\kappa\lambda\mu} = \tfrac{1}{2}(\partial_\lambda g_{\mu\kappa} + \partial_\mu g_{\lambda\kappa} - \partial_\kappa g_{\lambda\mu}).} \tag{7.3–12}$$

Thus we see that the Christoffel symbols of the first kind can be expressed in terms of the components of the metric tensor and are, therefore, isometric invariants. The same is true for the Christoffel symbols of the second kind, as follows at once from the second equation (7.3–7).

7.3.3 – The Christoffel symbols of the third kind

First we recall that the Christoffel symbols of the second kind are symmetric in the lower indices. This is a direct consequence of (7.3–6). Next we equate an upper and lower index and perform the summation according to the summation convention. We then get the expressions

$$\Gamma^\kappa{}_{\kappa\mu}, \tag{7.3–13}$$

referred to as ***the Christoffel symbols of the third kind.*** It is an easy matter to express these symbols in terms of the components of the metric tensor. This may be done by differentiating the determinant

$$g = \det\,[g_{\lambda\mu}].$$

To this end we observe that $gg^{\lambda\mu}$ is the cofacter $\dfrac{\partial g}{\partial g_{\lambda\tau}}$ of $g_{\lambda\mu}$. Hence, taking account of (7.3–9),

$$\partial_\mu g = \frac{\partial g}{\partial g_{\kappa\lambda}} \partial_\mu g_{\kappa\lambda} = g\, g^{\kappa\lambda} \partial_\mu g_{\kappa\lambda} = g\, g^{\kappa\lambda} \Gamma_{\kappa\lambda\mu} + g\, g^{\kappa\lambda} \Gamma_{\lambda\kappa\mu}$$
$$= g\,\Gamma^{\lambda}{}_{\lambda\mu} + g\,\Gamma^{\kappa}{}_{\kappa\mu} = 2g\,\Gamma^{\kappa}{}_{\kappa\mu}.$$

Thus we see that (observing that $g > 0$)

$$\boxed{\Gamma^{\kappa}{}_{\kappa\mu} = \tfrac{1}{2}\frac{\partial_\kappa g}{g} = \tfrac{1}{2}\partial_\kappa \log g = \partial_\kappa \log \sqrt{g} = \frac{\partial_\kappa \sqrt{g}}{\sqrt{g}}.} \qquad (7.3\text{–}14)$$

7.4 – Laws of transformation for the Christoffel symbols

7.4.1 – The symbols of the first kind

The Christoffel symbols do not transform like the components of a tensor. In order to obtain the laws of transformation we start from

$$\partial_{\lambda'} \boldsymbol{x} = \partial_\lambda \boldsymbol{x}\, \partial_{\lambda'} q^\lambda$$

and on differentiating with respect to $q^{\mu'}$ we find

$$\partial_{\lambda'\mu'} \boldsymbol{x} = \partial_{\lambda\mu} \boldsymbol{x}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + \boldsymbol{x}_\nu\, \partial_{\lambda'\mu'} q^\nu. \qquad (7.4\text{–}1)$$

On applying the operator P we get, in view of the first equation (7.3–4)

$$\boldsymbol{x}^{\kappa'} \Gamma_{\kappa'\lambda'\mu'} = \boldsymbol{x}^\kappa \Gamma_{\kappa\lambda\mu}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + \boldsymbol{x}_\nu\, \partial_{\lambda'\mu'} q^\nu,$$

whence, expressing $\boldsymbol{x}^{\kappa'}$ and $\boldsymbol{x}_\nu$ in terms of $\boldsymbol{x}^\kappa$,

$$\Gamma_{\kappa'\lambda'\mu'}\, \partial_\kappa q^{\kappa'} = \Gamma_{\kappa\lambda\mu}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + g_{\kappa\nu}\, \partial_{\lambda'\mu'} q^\nu, \qquad (7.4\text{–}2)$$

the desired equation of transformation for the Christoffel symbols of the first kind.

7.4.2 – The symbols of the second kind

On applying the operator P to (7.4–1) again and taking account of the second equation (7.3–4) we get (replacing in the last term ν by κ)

$$\boldsymbol{x}_{\kappa'} \Gamma^{\kappa'}{}_{\lambda'\mu'} = \boldsymbol{x}_\kappa \Gamma^\kappa{}_{\lambda\mu}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + \boldsymbol{x}_\kappa\, \partial_{\lambda'\mu'} q^\kappa,$$

whence, expressing $\boldsymbol{x}_{\kappa'}$ in terms of $\boldsymbol{x}_\kappa$,

$$\boxed{\Gamma^{\kappa'}{}_{\lambda'\mu'}\, \partial_{\kappa'} q^\kappa = \Gamma^\kappa{}_{\lambda\mu}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + \partial_{\lambda'\mu'} q^\kappa.} \qquad (7.4\text{–}3)$$

It should be noticed that this law of transformation does not involve explicitly the components of the metric tensor, in constrast with (7.4–2).

7.4.3 – The symbols of the third kind

By virtue of (5.4–14) the determinant $g = \det [g_{\lambda\mu}]$ transforms according to

$$\overset{(\kappa')}{g} = \overset{(\kappa)}{g}\, \Delta^{-2}, \tag{7.4–4}$$

where Δ is the jacobian

$$\frac{\partial(q^{1'}, \ldots, q^{n'})}{\partial(q^1, \ldots, q^n)}.$$

When differentiating both members of (7.4–4) logarithmically we obtain (omitting the superscript between small parenthesis)

$$\frac{\partial_{\mu'} g}{g} = \frac{\partial_\mu g}{g}\, \partial_{\mu'} q^\mu - 2\, \frac{\partial_\mu \Delta}{\Delta},$$

where $\partial_{\mu'} \Delta$ stands for $\partial_\mu \Delta\, \partial_{\mu'} q^\mu$. With reference to (7.3–14) we may infer that

$$\Gamma^{\kappa'}{}_{\kappa'\mu'} = \Gamma^\kappa{}_{\kappa\mu}\, \partial_{\mu'} q^\mu - \partial_{\mu'} \log |\Delta|. \tag{7.4–5}$$

7.5 – The components of the geodesic derivative

7.5.1 – The contravariant components

Let $\boldsymbol{v}(q)$ denote a tangential vector field along a curve $q^\kappa(q)$, $\kappa = 1, \ldots, n$. Differentiating the equation

$$\boldsymbol{v} = \boldsymbol{x}_\lambda v^\lambda \tag{7.5–1}$$

geodesically we have, in view of (7.1–4),

$$\mathsf{D}_q \boldsymbol{v} = (\mathsf{D}_q \boldsymbol{x}_\lambda)\, v^\lambda + \boldsymbol{x}_\kappa \mathrm{d}_q v^\kappa. \tag{7.5–2}$$

Next we observe that

$$\mathsf{D}_q \boldsymbol{x}_\lambda = \mathsf{P}\mathrm{d}_p \boldsymbol{x}_\lambda = \mathsf{P}\partial_{\lambda\mu} \boldsymbol{x}\, \mathrm{d}_q q^\mu = \boldsymbol{x}_\kappa \Gamma^\kappa{}_{\lambda\mu} \mathrm{d}_q q^\mu \tag{7.5–3}$$

and hence

$$\mathsf{D}_q \boldsymbol{v} = \boldsymbol{x}_\kappa (\mathrm{d}_q v^\kappa + \Gamma^\kappa{}_{\lambda\mu} v^\lambda \mathrm{d}_q q^\mu).$$

It is customary to write

$$\mathsf{D}_q \boldsymbol{v} = \boldsymbol{x}_\kappa \mathrm{D}_q v^\kappa, \tag{7.5–4}$$

and thus we have

$$\boxed{\mathrm{D}_q v^\kappa = \mathrm{d}_q v^\kappa + \Gamma^\kappa{}_{\lambda\mu} v^\lambda \mathrm{d}_q q^\mu.} \tag{7.5–5}$$

These expressions are called the *geodesic derivatives of the contravariant components of the vector* $\boldsymbol{v}$. It appears that *the process of geodesic differentiation yields again components of a vector.*

7.5.2 – The covariant components

The expression of $\boldsymbol{v}$ in terms of covariant components is

$$\boldsymbol{v} = \boldsymbol{x}^\lambda v_\lambda, \tag{7.5–6}$$

whence

$$\mathsf{D}_q \boldsymbol{v} = (\mathsf{D}_q \boldsymbol{x}^\kappa) v_\kappa + \boldsymbol{x}^\lambda \mathrm{d}_q v_\lambda. \tag{7.5–7}$$

Next we wish to obtain an expression for $\mathsf{D}_q \boldsymbol{x}^\kappa$ similar to (7.5–3). From

$$\boldsymbol{x}^\kappa \boldsymbol{x}_\lambda = \delta^\kappa_\lambda$$

follows

$$(\mathsf{D}_q \boldsymbol{x}^\kappa)\, \boldsymbol{x}_\lambda + \boldsymbol{x}^\kappa\, (\mathsf{D}_q \boldsymbol{x}_\lambda) = 0,$$

or, in view of (7.5–2),

$$\boldsymbol{x}_\lambda\, (\mathsf{D}_q \boldsymbol{x}^\kappa) = -\boldsymbol{x}^\kappa \boldsymbol{x}_\nu \Gamma^\nu_{\lambda\mu} \mathrm{d}_q q^\mu = -\Gamma^\kappa{}_{\lambda\mu} \mathrm{d}_q q^\mu.$$

As a consequence we have

$$\mathsf{D}_q \boldsymbol{x}^\kappa = -\boldsymbol{x}^\lambda \Gamma^\kappa{}_{\lambda\mu} \mathrm{d}_q q^\mu. \tag{7.5–8}$$

Thus we derive from (7.5–6)

$$\mathsf{D}_q \boldsymbol{v} = \boldsymbol{x}^\lambda (\mathrm{d}_q v_\lambda - \Gamma^\kappa{}_{\lambda\mu} v_\kappa \mathrm{d}_q q^\mu).$$

If we agree to write

$$\mathsf{D}_q \boldsymbol{v} = \boldsymbol{x}^\lambda \mathrm{D}_q v_\lambda \tag{7.5–9}$$

we also have

$$\boxed{\mathrm{D}_q v_\lambda = \mathrm{d}_q v_\lambda - \Gamma^\kappa{}_{\lambda\mu} v_\kappa \mathrm{d}_q q^\mu,} \tag{7.5–10}$$

the *geodesic derivatives of the covariant components of the vectors* $\boldsymbol{v}$.

In view of (7.5–5) and (7.5–10) it is clear now, that *geodesic differentiation is an isometrically invariant process.*

As a by-product of (7.5–8) we have

$$\mathsf{P}\partial_\mu \boldsymbol{x}^\kappa = -\boldsymbol{x}^\lambda \Gamma^\kappa{}_{\lambda\mu}. \tag{7.5–11}$$

It is an easy matter to verify this result without the application of covariant differentiation.

7.5.3 – Lowering and raising the index

We wish to compare the two expressions (7.5–4) and (7.5–9) for the geodesic derivative of the vector $\boldsymbol{v}$. In

$$\boldsymbol{x}_\lambda \mathrm{D}_q v^\lambda = \boldsymbol{x}^\mu \mathrm{D}_q v_\mu$$

we may express $\boldsymbol{x}_\lambda$ in terms of $\boldsymbol{x}^\mu$ and we find

$$g_{\lambda\mu} \mathrm{D}_q v^\lambda = \mathrm{D}_q v_\mu,$$

whence

$$\boxed{g_{\lambda\mu} \mathrm{D}_q v^\lambda = \mathrm{D}_q (g_{\lambda\mu} v^\lambda).} \tag{7.5–12}$$

On expressing $\boldsymbol{x}^\mu$ in terms of $\boldsymbol{x}_\lambda$ we find in exactly the same way

$$\boxed{g^{\lambda\mu} \mathrm{D}_q v_\mu = \mathrm{D}_q (g^{\lambda\mu} v_\mu).} \tag{7.5–13}$$

Hence *the process of lowering or raising the index and finding the geodesic derivative are interchangeable.*

7.5.4 – The differential equations of the geodesic curves

We may apply (7.5–5) to the case that the field $\boldsymbol{v}(q)$ coincides with the field $\mathrm{d}_q \boldsymbol{x}$ of derivatives of $\boldsymbol{x}(q)$. Then we find that the geodesic curves are characterized by the system of differential equations

$$\mathrm{d}_q{}^2 q^\kappa + \Gamma^\kappa{}_{\lambda\mu} \mathrm{d}_q q^\lambda \mathrm{d}_q q^\mu = \alpha \mathrm{d}_q q^\kappa, \qquad \kappa = 1, \ldots, n, \tag{7.5–14}$$

where α is a function of the parameter q. If the curve is referred to a natural parameter the system takes the simpler form

$$\boxed{\mathrm{d}_t{}^2 q^\kappa + \Gamma^\kappa{}_{\lambda\mu} \mathrm{d}_t q^\lambda \mathrm{d}_t q^\mu = 0.} \tag{7.5–15}$$

Thus it appears that the geodesic curves are obtained by solving a system of ordinary differential equations of the second order. It follows from the existence theorem for equations of this kind, that there is always one and only one set of functions $q^\kappa(t)$ satisfying (7.5–15) such that these functions and their first derivatives take prescribed values for a given value of t. We may state this in the following way:

Through a given point on a manifold there is one and only one geodesic curve having there a prescribed direction.

7.5.5 – An extremal property of geodesic curves

Let $\varphi(q^1, \ldots, q^n)$ denote a scalar invariant defined throughout a certain region on an n-dimensional manifold. The set of points at which φ takes a fixed value is called a *level-manifold of* φ. Let us assume that at all points of the level-manifold $\varphi = 0$ the gradient does not vanish (within a certain region). Then through each point of this submanifold passes exactly one geodesic curve having the direction grad φ as tangent.

By introducing a new system of parameters such that $q^1 = \varphi$, the geodesic curves become parameter curves. In particular

$$g_{1k} = 0, \qquad k = 2, \ldots, n, \tag{7.5–16}$$

at the points of $\varphi = 0$.

The arc length s along a geodesic measured from $\varphi = 0$ defines another scalar invariant. By transforming q^1 again it is possible to identify s and q^1 and it follows that

$$g_{11} = 1, \tag{7.5–17}$$

for along the q^1-curves we have

$$1 = (\mathrm{d}_s s)^2 = g_{11}\,\mathrm{d}_s q^1\,\mathrm{d}_s q^1 = g_{11}.$$

From the differential equations, satisfied by the geodesics, viz.

$$\mathrm{d}_s{}^2 q^k + \Gamma^k{}_{lm}\,\mathrm{d}_s q^l\,\mathrm{d}_s q^m = 0,$$

we find, by observing that $s = q^1$:

$$\Gamma^k{}_{11} = 0, \qquad k = 1, \ldots, n,$$

and on account of (7.3–7)

$$\Gamma_{k11} = 0.$$

In view of (7.5–17) we evidently have

$$\partial_1\, g_{1k} = 0, \qquad k = 1, \ldots, n$$

and therefore the equations (7.5–16) hold everywhere, i.e., the directions of the geodesics are given by grad φ.

Let us now consider a curve $q^\kappa(s)$ connecting two points $s = 0$ and $s = a$ on the same geodesic. Here s is not necessarily the arc length of the curve. The length of this arc is

$$\begin{aligned}\hat{a} &= \int_0^a \sqrt{\mathrm{d}_s q^1\,\mathrm{d}_s q^1 + g_{hk}\,\mathrm{d}_s q^h\,\mathrm{d}_s q^k} \\ &= \int_0^a \sqrt{1 + g_{hk}\,\mathrm{d}_s q^h\,\mathrm{d}_s q^k}\;\mathrm{d}s \geqq \int_0^a \mathrm{d}s = a,\end{aligned}$$

where $h, k = 2, \ldots, n$. The system of geodesics is such, that at every point their directions coincide with the gradient of a scalar invariant. Such a system is called a *normal congruence*. Thus we have proved:

When a geodesic curve can be imbedded in a normal congruence of geodesics the distance of two points on it, measured by the arc length, is not greater than the length of any other arc connecting these points.

This theorem generalizes the well-known property of a linear segment connecting two points in a flat space.

In section 6.3.1 we already encountered an example of a normal congruence of geodesics, viz. the generators of the surface of tangents of a space curve. In this case the level curves are the involutes.

7.6 – Geodesic and conformal correspondences

7.6.1 – Geodesic correspondence

An interesting problem arises when we consider two manifolds determined by the vector functions $\boldsymbol{x}(q^\kappa)$ and $\overset{*}{\boldsymbol{x}}(q^\kappa)$ which correspond in such a way that when a system of functions $q^\kappa(q)$ assigns a geodesic curve on one of the manifolds, then it also assigns a geodesic curve on the other manifold. This correspondence is called a *geodesic correspondence* and may be considered as a counter part of a projective correspondence between linear spaces. An example is provided by an isometric correspondence, for geodesic curves are isometrically invariant. In that case the Christoffel symbols $\Gamma^\kappa{}_{\lambda\mu}$ and $\Gamma^{*\kappa}{}_{\lambda\mu}$ evaluated at corresponding points are equal.

We wish to find the relation between corresponding Christoffel symbols when we are dealing with a general geodesic mapping.

When $q^\kappa(q)$, $\kappa = 1, \ldots, n$, determine a geodesic curve on either manifold, then there exist functions α and $\overset{*}{\alpha}$ such that

$$d_q{}^2 q^\kappa + \Gamma^\kappa{}_{\lambda\mu} d_q q^\lambda d_q q^\mu = \alpha d_q q^\kappa \tag{7.6–1}$$

and

$$d_q{}^2 q^\kappa + \overset{*}{\Gamma}{}^\kappa{}_{\lambda\mu} d_q q^\lambda d_q q^\mu = \overset{*}{\alpha} d_q q^\kappa, \tag{7.6–2}$$

these equations being the differential equations (7.5–13). On subtracting corresponding members of these equations we get

$$\Phi^\kappa{}_{\lambda\mu} d_q q^\lambda d_q q^\mu = (\overset{*}{\alpha} - \alpha) d_q q^\kappa, \qquad \kappa = 1, \ldots, n, \tag{7.6–3}$$

with

$$\Phi^\kappa{}_{\lambda\mu} = \overset{*}{\Gamma}{}^\kappa{}_{\lambda\mu} - \Gamma^\kappa{}_{\lambda\mu}. \tag{7.6–4}$$

First we observe that the expressions $\Phi^{\kappa}{}_{\lambda\mu} = \Phi^{\kappa}{}_{\mu\lambda}$ are the components of a tensor. In fact, in addition to (7.4–3) we also have

$$\overset{*}{\Gamma}{}^{\kappa'}{}_{\lambda'\mu'}\partial_{\kappa'}q^{\kappa} = \overset{*}{\Gamma}{}^{\kappa}{}_{\lambda\mu}\partial_{\lambda'}q^{\lambda}\partial_{\mu'}q^{\mu} + \partial_{\lambda'\mu'}q^{\kappa}, \tag{7.6–5}$$

and on subtracting corresponding members of (7.6–5) and (7.4–3) we get

$$\Phi^{\kappa'}{}_{\lambda'\mu'}\partial_{\kappa'}q^{\kappa} = \Phi^{\kappa}{}_{\lambda\mu}\partial_{\lambda'}q^{\lambda}\partial_{\mu'}q^{\mu}$$

which proves the assertion.

Eliminating $(\overset{*}{\alpha} - \alpha)$ from the system (7.6–3) we arrive at

$$\Phi^{\kappa}{}_{\lambda\mu}\,\mathrm{d}_q q^{\lambda}\,\mathrm{d}_q q^{\mu}\,\mathrm{d}_q q^{\nu} = \Phi^{\nu}{}_{\lambda\mu}\,\mathrm{d}_q q^{\lambda}\,\mathrm{d}_q q^{\mu}\,\mathrm{d}_q q^{\kappa}.$$

It is more convenient to write these equations in the form

$$\delta^{\nu}_{\rho}\Phi^{\kappa}{}_{\lambda\mu}\,\mathrm{d}_q q^{\lambda}\,\mathrm{d}_q q^{\mu}\,\mathrm{d}_q q^{\rho} = \delta^{\kappa}_{\rho}\Phi^{\nu}{}_{\lambda\mu}\,\mathrm{d}_q q^{\lambda}\,\mathrm{d}_q q^{\mu}\,\mathrm{d}_q q^{\rho}.$$

These equations are satisfied identically in $\mathrm{d}_q q^{\kappa}$, $\kappa = 1 \ldots, n$, and these derivatives may be chosen in an arbitrary way. Hence

$$\delta^{\nu}_{\rho}\Phi^{\kappa}{}_{\lambda\mu} + \delta^{\nu}_{\lambda}\Phi^{\kappa}{}_{\mu\rho} + \delta^{\nu}_{\mu}\Phi^{\kappa}{}_{\rho\lambda} = \delta^{\kappa}_{\rho}\Phi^{\nu}{}_{\lambda\mu} + \delta^{\nu}_{\lambda}\Phi^{\nu}{}_{\mu\rho} + \delta^{\kappa}_{\mu}\Phi^{\nu}{}_{\rho\lambda}.$$

Contracting with respect to ν and ρ we get

$$n\Phi^{\kappa}{}_{\lambda\mu} + \Phi^{\kappa}{}_{\mu\lambda} + \Phi^{\kappa}{}_{\mu\lambda} = \Phi^{\kappa}{}_{\lambda\mu} + \delta^{\kappa}_{\lambda}\Phi^{\nu}{}_{\mu\nu} + \delta^{\kappa}_{\mu}\Phi^{\nu}{}_{\nu\lambda}$$

or

$$(n+1)\Phi^{\kappa}{}_{\lambda\mu} = \delta^{\kappa}_{\lambda}\Phi^{\nu}{}_{\nu\mu} + \delta^{\kappa}_{\mu}\Phi^{\nu}{}_{\nu\lambda}, \tag{7.6–6}$$

taking account of the symmetry of $\Phi^{\kappa}{}_{\lambda\mu}$ with respect to the lower indices.

Since the $\Phi^{\kappa}{}_{\lambda\mu}$ are the components of a tensor the expressions $\Phi^{\nu}{}_{\nu\mu}$ and $\Phi^{\nu}{}_{\nu\lambda}$ occurring on the right of (7.6–6) are the components of a vector. If we introduce the components φ_{κ} determined by

$$(n+1)\varphi_{\kappa} = \Phi^{\nu}{}_{\nu\kappa} \tag{7.6–7}$$

then (7.6–6) appears in the form

$$\Phi^{\kappa}{}_{\lambda\mu} = \delta^{\kappa}_{\lambda}\varphi_{\mu} + \delta^{\kappa}_{\mu}\varphi_{\lambda}. \tag{7.6–8}$$

Thus we have proved the first half of the following theorem:

A necessary and sufficient condition for the manifolds $\boldsymbol{x}(q^{\kappa})$ *and* $\overset{*}{\boldsymbol{x}}(q^{\kappa})$ *to be in geodesic correspondence is expressed by the relations*

$$\boxed{\overset{*}{\Gamma}{}^{\kappa}{}_{\lambda\mu} = \Gamma^{\kappa}{}_{\lambda\mu} + \delta^{\kappa}_{\lambda}\varphi_{\mu} + \delta^{\kappa}_{\mu}\varphi_{\lambda},} \tag{7.6–9}$$

where φ_{κ} *are the components of a vector.*

Suppose that the conditions (7.6–9) are satisfied and consider a geodesic curve on the first manifold characterized by the equations (7.6–1). Next we observe that

$$(\delta^\kappa_\lambda \varphi_\mu + \delta^\kappa_\mu \varphi_\lambda)\, \mathrm{d}_a q^\lambda\, \mathrm{d}_a q^\mu = 2\varphi_\nu\, \mathrm{d}_a q^\nu\, \mathrm{d}_a q^\kappa, \tag{7.6–10}$$

Since $\varphi_\nu \mathrm{d}_a q^\nu$ is a scalar, we may introduce the function

$$\overset{*}{\alpha} = \alpha + 2\varphi_\nu\, \mathrm{d}_a q^\nu \tag{7.6–11}$$

and, by adding the corresponding members of (7.6–1) and (7.6–2), we get the system of (7.6–2).

Additionally we have:

The vector φ_κ, $\kappa = 1, \ldots, n$, occurring in (7.6–9) is a gradient.

In fact, it follows from (7.6–7) and (7.6–4) that

$$(n+1)\varphi_\kappa = \overset{*}{\Gamma}{}^\nu{}_{\nu\kappa} - \Gamma^\nu{}_{\nu\kappa}$$

and from (7.4–5) we may conclude that

$$(n+1)\varphi_\kappa = \tfrac{1}{2}\partial_\kappa \log \overset{*}{g} - \tfrac{1}{2}\partial_\kappa \log g = \partial_\kappa \log \sqrt{\overset{*}{g}/g}, \tag{7.6–12}$$

where $\overset{*}{g}$ denotes the determinant of the metric tensor on the second manifold. This proves the assertion.

7.6.2 – The Thomas symbols for geodesic correspondence

It is apparent from (7.6–9) that the Christoffel symbols are not invariant with respect to geodesic correspondence. It is not difficult, however, to obtain expressions which remain invariant under geodesic mapping. The equation (7.6–8) and (7.6–7) may be combined into

$$\Phi^\kappa{}_{\lambda\mu} = \frac{1}{n+1}\,\delta^\kappa_\lambda\, \Phi^\nu{}_{\nu\mu} + \frac{1}{n+1}\,\delta^\kappa_\mu\, \Phi^\nu{}_{\nu\lambda}$$

and with reference to (7.6–4) we may conclude that

$$\overset{*}{\Pi}{}^\kappa{}_{\lambda\mu} = \Pi^\kappa{}_{\lambda\mu}, \tag{7.6–13}$$

whereby

$$\boxed{\Pi^\kappa{}_{\lambda\mu} = \Gamma^\kappa{}_{\lambda\mu} - \frac{1}{n+1}\,\delta^\kappa_\lambda\, \Gamma^\nu{}_{\nu\mu} - \frac{1}{n+1}\,\delta^\kappa_\mu\, \Gamma^\nu{}_{\nu\lambda}} \tag{7.6–14}$$

are the desired invariants. They are called the *projective Thomas symbols.*

Making use of these symbols we can write the equation (7.6–1) as

$$\mathrm{d}_q{}^2 q^\kappa + \Pi^\kappa{}_{\lambda\mu}\, \mathrm{d}_q q\, \mathrm{d}_q{}^\lambda q^\mu = \beta\, \mathrm{d}_q q^\kappa, \tag{7.6–15}$$

with

$$\beta = \alpha - \frac{2}{n+1}\, \Gamma^\nu{}_{\nu\kappa}\, \mathrm{d}_q q^\kappa = \alpha - \partial_\kappa \log g\, \mathrm{d}_q q^\kappa.$$

It follows that $\overset{*}{\beta} = \beta$, where $\overset{*}{\beta}$ is the analogous expression derived from the equation (7.6–2). But this is also clear from (7.6–11) and (7.6–12).

Reasoning along the same lines as in section 7.6.2 we may establish the existence of a parameter ϑ such that the equation (7.6–15) takes the simpler form

$$\mathrm{d}_\vartheta{}^2 q^\kappa + \Pi^\kappa{}_{\lambda\mu}\, \mathrm{d}_\vartheta q^\lambda\, \mathrm{d}_\vartheta q^\mu = 0, \tag{7.6–16}$$

where ϑ is a solution of the differential equation

$$\mathrm{d}_q{}^2 \vartheta = \beta\, \mathrm{d}_q \vartheta.$$

7.6.3 – Conformal correspondence

The condition for a conformal correspondence between the manifolds $\boldsymbol{x}(q^\kappa)$ and $\overset{*}{\boldsymbol{x}}(q^\kappa)$ is expressed by (5.3–18). As a consequence the relation between the determinants g and $\overset{*}{g}$ of the respective metric tensors is

$$\overset{*}{g} = \lambda^n g. \tag{7.6–17}$$

It follows that λ^n is positive and without loss of generality we may assume that λ is positive, too. For our purpose it is more convenient to write

$$\lambda = \mathrm{e}^{2\chi}, \tag{7.6–18}$$

where χ is a function of q^κ, $\kappa = 1, \ldots, n$, and e the basis of natural logarithms. Thus we have

$$\overset{*}{g}_{\lambda\mu} = \mathrm{e}^{2\chi} g_{\lambda\mu} \tag{7.6–19}$$

and, consequently,

$$\overset{*}{g}{}^{\lambda\mu} = \mathrm{e}^{-2\chi} g^{\lambda\mu}, \tag{7.6–20}$$

where $\overset{*}{g}{}^{\lambda\mu}$ is such that $\overset{*}{g}{}^{\lambda\kappa}\overset{*}{g}_{\kappa\mu} = \delta^\lambda_\mu$. In this connection it should be noticed that the $\overset{*}{g}_{\lambda\mu}$ may be considered as the components of a tensor defined on the manifold $\boldsymbol{x}(q^\kappa)$. But the $\overset{*}{g}{}^{\lambda\mu}$ are not obtained by applying the process of raising the indices by means of $g^{\lambda\mu}$. The danger of confusion is, however, remote.

Next we wish to establish a relation between the corresponding Christoffel symbols. An easy computation based on (7.3–12) leads to

$$\mathrm{e}^{-2\chi}\overset{*}{\Gamma}_{\kappa\lambda\mu} = \Gamma_{\kappa\lambda\mu}+g_{\mu\kappa}\partial_\lambda\chi+g_{\kappa\lambda}\partial_\mu\chi-g_{\lambda\mu}\partial_\nu\chi. \tag{7.6–21}$$

Taking account of (7.6–20) we readily find

$$\boxed{\overset{*}{\Gamma}{}^{\kappa}{}_{\lambda\mu} = \Gamma^{\kappa}{}_{\lambda\mu}+\delta^{\kappa}_{\mu}\partial_\lambda\chi+\delta^{\kappa}_{\lambda}\partial_\mu\chi-g^{\kappa\nu}g_{\lambda\mu}\partial_\nu\chi,} \tag{7.6–22}$$

the desired relation between Christoffel symbols of the second kind. The difference between corresponding symbols is, therefore, the tensor

$$\Phi^{\nu}{}_{\lambda\mu} = \delta^{\kappa}_{\lambda}\partial_\mu\chi+\delta^{\kappa}_{\mu}\partial_\lambda\chi-g^{\kappa\nu}g_{\lambda\mu}\partial_\nu\chi. \tag{7.6–23}$$

7.6.4 – The conformal Thomas symbols

Just as in section 7.6.2 we may now derive conformal invariants of the type of Christoffel symbols. Contraction with respect to κ and λ yields from (7.6–23)

$$\Phi^{\nu}{}_{\nu\mu} = n\partial_\mu\chi+\delta^{\nu}_{\mu}\partial_\mu\chi-\delta^{\nu}_{\mu}\partial_\nu\chi = n\partial_\mu\chi, \tag{7.6–24}$$

and thus (7.6–23) appears as

$$\Phi^{\kappa}{}_{\lambda\mu} = \frac{1}{n}\delta^{\kappa}_{\lambda}\Phi^{\nu}{}_{\nu\mu}+\frac{1}{n}\delta^{\kappa}_{\mu}\Phi^{\nu}{}_{\nu\lambda}-\frac{1}{n}g^{\kappa\nu}g_{\lambda\mu}\Phi^{\rho}{}_{\rho\nu}. \tag{7.6–25}$$

This is equivalent to

$$\overset{*}{K}{}^{\kappa}{}_{\lambda\mu} = K^{\kappa}{}_{\lambda\mu}, \tag{7.6–26}$$

where

$$\boxed{K^{\kappa}{}_{\lambda\mu} = \Gamma^{\kappa}{}_{\lambda\mu}-\frac{1}{n}\delta^{\kappa}_{\lambda}\Gamma^{\nu}{}_{\nu\mu}-\frac{1}{n}\delta^{\kappa}_{\mu}\Gamma^{\nu}{}_{\nu\lambda}+\frac{1}{n}g^{\kappa\nu}g_{\lambda\mu}\Gamma^{\rho}{}_{\rho\nu}} \tag{7.6–27}$$

are the *conformal Thomas symbols.*

By making use of (7.3–7), (7.3–12) and (7.3–14) we can exhibit a remarkable analogy between these Thomas symbols and the Christoffel symbols expressed in terms of the components of the metric tensor. In fact, we may write (7.6–27) as

$$\begin{aligned} K^{\kappa}{}_{\lambda\mu} &= g^{\kappa\nu}\,\Gamma_{\nu\lambda\mu}-\frac{1}{2n}\delta^{\kappa}_{\lambda}\frac{\partial_\mu g}{g}-\frac{1}{2n}\delta^{\kappa}_{\mu}\frac{\partial_\lambda g}{g}+\frac{1}{2n}g^{\kappa\nu}g_{\lambda\mu}\frac{\partial_\nu g}{g} \\ &= \tfrac{1}{2}g^{\kappa\nu}\left(\partial_\lambda g_{\mu\nu}-\frac{1}{n}g_{\mu\nu}\frac{\partial_\lambda g}{g}+\partial_\mu g_{\nu\lambda}-\frac{1}{n}g_{\nu\lambda}\frac{\partial_\mu g}{g}-\partial_\nu g_{\lambda\mu}+\frac{1}{n}g_{\lambda\mu}\frac{\partial_\nu g}{g}\right). \end{aligned}$$

Introducing the modified components of the metric tensor

$$\tilde{g}_{\lambda\mu} = g_{\lambda\mu} g^{-1/n}, \qquad \tilde{g}^{\lambda\mu} = g^{\lambda\mu} g^{1/n}, \tag{7.6–28}$$

which are evidently conformal invariants, we readily get

$$K^{\kappa}{}_{\lambda\mu} = \tfrac{1}{2}\tilde{g}^{\kappa\nu}(\partial_\lambda \tilde{g}_{\mu\nu} + \partial_\mu \tilde{g}_{\nu\lambda} - \partial_\nu \tilde{g}_{\lambda\mu}). \tag{7.6–29}$$

It should be noticed that the expressions (7.6–28) are not the components of a tensor, according to the fact that g is a density.

Finally we remark that contraction of the projective and conformal Thomas symbols does not yield new invariants, for the contracted symbols are zero.

7.6.5 – Conformal geodesic correspondence

We wish to give an answer to the following question. Under what circumstances is a conformal correspondence between two manifolds also a geodesic mapping? This problem may be solved by equating the two expressions (7.6–8) and (7.6–23):

$$\delta^\kappa_\lambda \varphi_\mu + \delta^\kappa_\mu \varphi_\lambda = \delta^\kappa_\lambda \partial_\mu \chi + \delta^\kappa_\mu \partial_\lambda \chi - g^{\kappa\nu} g_{\lambda\mu} \partial_\nu \chi. \tag{7.6–30}$$

Contraction with respect to κ and λ yields

$$(n+1)\varphi_\mu = (n+1)\partial_\mu \chi - \partial_\mu \chi = n\partial_\mu \chi$$

or

$$\varphi_\mu = \frac{n}{n+1}\partial_\mu \chi. \tag{7.6–31}$$

Inserting this into (7.6–30) we get

$$\delta^\kappa_\lambda \partial_\mu \chi + \delta^\kappa_\mu \partial_\lambda \chi = (n+1) g^{\kappa\nu} g_{\lambda\mu} \partial_\nu \chi. \tag{7.6–32}$$

Transvection by $g^{\lambda\mu}$ gives

$$g^{\kappa\mu}\partial_\mu \chi + g^{\kappa\lambda}\partial_\lambda \chi = n(n+1) g^{\kappa\nu}\partial_\nu \chi, \tag{7.6–33}$$

whence

$$2\partial_\kappa \chi = n(n+1)\partial_\kappa \chi$$

or

$$(n-1)(n+2)\partial_\kappa \chi = 0. \tag{7.6–34}$$

Hence, since $n \geqq 2$, $\partial_\kappa \chi = 0$, $\kappa = 1, \ldots, n$, and it follows that χ is a constant. Thus we have found:

A conformal correspondence is also a geodesic correspondence if and only if the function χ in (7.6–19) is a constant.

7.7 – Ricci's coefficients of rotation

7.7.1 – Infinitesimal rotations

In an n-dimensional metric vector space we consider an orthonormal frame $\underset{h}{e}$, $h = 1, \ldots, n$, whose elements are differentiable functions of a variable q. The derivatives of these vectors evaluated for a given value of q may be referred to the frame corresponding to the same value of q. We then have

$$\mathrm{d}_q \underset{k}{e} = \sum_{h=1}^{n} \underset{h}{e} \underset{hk}{f}, \qquad k = 1, \ldots, n. \tag{7.7–1}$$

It follows that

$$\underset{hk}{f} = \underset{h}{e}\, \mathrm{d}_q \underset{k}{e}. \tag{7.7–2}$$

Since $\underset{h}{e}\,\underset{k}{e}$ is either zero or one, we also have

$$\underset{h}{e}\, \mathrm{d}_q \underset{k}{e} + \underset{k}{e}\, \mathrm{d}_q \underset{h}{e} = 0.$$

Hence

$$\underset{hk}{f} = -\underset{kh}{f}. \tag{7.7–3}$$

The linear transformation (7.7–1) turns out to be anti-symmetric. This transformation is called an *infinitesimal rotation* and the coefficients $\underset{hk}{f}$ are the *coefficients of rotation*. It is our aim to study entities defined on an arbitrary manifold wh·ch show a close similarity to these coefficients.

7.7.2 – Congruences of curves

A system of curves on a manifold such that through an arbitrary point just one curve passes, is called a *congruence*. An example of a congruence is a system of parameter curves. They are obtained when we keep fixed all variables except one in the function $x(q^1, \ldots, q^n)$. Thus we can assign n congruences of parameter curves on the manifold.

We wish to study an *orthogonal system* of n congruences. The curves of each congruence are assumed to be referred to the arc length as parameter. The tangent vectors $\underset{h}{e}$, $h = 1, \ldots, n$, of the curves passing through a given point at this point are assumed to constitute an orthonormal set.

Next we consider an arbitrary curve $q^\kappa(s)$, $\kappa = 1, \ldots, n$, also referred to its arc length. The geodesic derivative of $\underset{h}{e}$ at a given point with respect to this curve referred to the above mentioned frame at this point may be written as

$$\mathsf{D}_s \underset{k}{\boldsymbol{e}} = \sum_{h=1}^{n} \underset{h}{\boldsymbol{e}} \underset{hk}{f}, \tag{7.7-4}$$

with

$$\underset{hk}{f} = \underset{h}{\boldsymbol{e}} \, \mathsf{D}_s \underset{k}{\boldsymbol{e}}, \tag{7.7-5}$$

and it follows as in the previous section that the coefficients (7.7–5) are anti-symmetric. They are called the *coefficients of rotation* of the orthogonal system of congruences with respect to the given curve.

7.7.3 – Ricci's coefficients of rotation

The congruences of an orthonormal set will be labelled by the tangent vectors of their curves. We are particularly interested in the coefficients of rotation of the system $\underset{l}{\boldsymbol{e}}$ with respect to the curves belonging to any one of the congruences $\underset{m}{\boldsymbol{e}}$ of the system. In other words, we wish to study the geodesic derivatives of the vectors $\underset{l}{\boldsymbol{e}}$ with respect to a curve belonging to the congruence $\underset{m}{\boldsymbol{e}}$, that is to say, the congruence whose curves yield the tangent field $\underset{m}{\boldsymbol{e}}$. This derivative will be denoted by $\underset{lm}{\boldsymbol{e}}$. Referred to a frame afforded by the tangent vectors at a given point, this vector may be represented by

$$\underset{lm}{\boldsymbol{e}} = \sum_{k=1}^{n} \underset{k}{\boldsymbol{e}} \underset{klm}{\gamma}, \qquad l, m = 1, \ldots, n, \tag{7.7-6}$$

with

$$\underset{klm}{\gamma} = \underset{k}{\boldsymbol{e}} \, \underset{lm}{\boldsymbol{e}}. \tag{7.7-7}$$

These coefficients were introduced by Ricci.

Since

$$\underset{k}{\boldsymbol{e}} \, \underset{lm}{\boldsymbol{e}} + \underset{l}{\boldsymbol{e}} \, \underset{km}{\boldsymbol{e}} = 0$$

we have

$$\boxed{\underset{klm}{\gamma} + \underset{lkm}{\gamma} = 0,} \qquad k, l, m = 1, \ldots, n, \tag{7.7-8}$$

expressing the anti-symmetry with respect to the first pair of subscripts.

The Ricci coefficients are closely related to the Christoffel three index symbols of the second kind. Let s_m denote the arc length of any curve of the congruence $\underset{m}{\boldsymbol{e}}$. In view of (7.5–5) the covariant components of $\underset{lm}{\boldsymbol{e}}$ with respect to an arbitrary frame (k) are

$$\mathrm{d}_{s_m} \underset{l}{e}^{\kappa} + \Gamma^{\kappa}{}_{\lambda\mu} \underset{l}{e}^{\lambda} \underset{m}{e}^{\mu}.$$

Hence

$$\underset{k}{\boldsymbol{e}}\,\underset{lm}{\boldsymbol{e}} = \underset{k}{e_\nu}\,\mathrm{d}_{s_m}\underset{l}{e^\nu} + \Gamma^\kappa{}_{\lambda\mu}\underset{l}{e^\lambda}\,\underset{m}{e^\mu}\,\underset{k}{e_\kappa}.$$

Thus we find

$$\boxed{\underset{klm}{\gamma} = \Gamma^\kappa{}_{\lambda\mu}\underset{k}{e_\kappa}\,\underset{l}{e_\lambda}\,\underset{m}{e_\mu} + \underset{k}{e_\nu}\,\mathrm{d}_{s_m}\underset{l}{e^\nu}.} \tag{7.7–9}$$

This formula shows a close resemblance with (7.4–3).

7.7.4 – Geodesic and normal congruences

A congruence is said to be *geodesic* if all its curves are geodesic curves. Suppose that the congruence $\underset{n}{\boldsymbol{e}}$ is geodesic. Then $\underset{nn}{\boldsymbol{e}} = 0$ and $\underset{k}{\boldsymbol{e}}\,\underset{nn}{\boldsymbol{e}} = 0$. Hence:

A necessary condition for the congruence $\underset{n}{\boldsymbol{e}}$ to be geodesic is expressed by

$$\boxed{\underset{k\mathrm{nn}}{\gamma} = 0,} \qquad k = 1, \ldots, \mathrm{n}. \tag{7.7–10}$$

It is clear that this condition is also sufficient.

A congruence is said to be *normal* if the tangent vectors of its curves are proportional to the gradient field of a scalar invariant, i.e., if there are functions φ and ψ such that

$$\partial_\lambda \varphi = \psi e_\lambda, \qquad \lambda = 1, \ldots, n, \tag{7.7–11}$$

where the e_λ denote the covariant components of the tangent vectors of the curves of the congruence. Differentiating again yields

$$\partial_{\lambda\mu}\varphi = e_\lambda \partial_\mu \psi + \psi \partial_\mu e_\lambda,$$

whence, on account of the symmetry of $\partial_{\lambda\mu}\varphi$,

$$e_\lambda \partial_\mu \psi + \psi \partial_\mu e_\lambda = e_\mu \partial_\lambda \psi + \psi \partial_\lambda e_\mu$$

or

$$e_\lambda \partial_\mu \psi - e_\mu \partial_\lambda \psi = \psi(\partial_\lambda e_\mu - \partial_\mu e_\lambda).$$

A direct consequence is the relation

$$(\partial_\lambda e_\mu - \partial_\mu e_\lambda)e_\kappa + (\partial_\mu e_\kappa - \partial_\kappa e_\mu)e_\lambda + (\partial_\lambda e_\mu - \partial_\mu e_\lambda)e_\kappa = 0, \tag{7.7–12}$$

The covariant components of the geodesic derivative of $\boldsymbol{e}$ with respect to a curve of the congruence $\underset{m}{\boldsymbol{e}}$ are

$$\mathrm{d}_{s_m} e_\kappa - \Gamma^\nu{}_{\kappa\mu} e_\nu \underset{m}{e^\mu} = (\partial_\mu e_\kappa - \Gamma^\nu{}_{\kappa\mu} e_\nu)\underset{m}{e^\mu}.$$

Applied to the curves of the congruence $\underset{n}{\boldsymbol{e}}$ we get

$$\underset{k\,n\,m}{\boldsymbol{e}\,\boldsymbol{e}} = (\partial_\mu \underset{n}{e_\kappa} - \Gamma^\nu{}_{\kappa\mu} \underset{n}{e_\nu}) \underset{k}{e^\kappa} \underset{m}{e^\mu}, \tag{7.7-13}$$

for

$$d_{s_m} \underset{n}{e_\kappa} = (\partial_\mu \underset{n}{e_\kappa}) \underset{m}{e^\mu}.$$

Hence

$$\underset{k\,nm}{\boldsymbol{e}\,\boldsymbol{e}} - \underset{m\,nk}{\boldsymbol{e}\,\boldsymbol{e}} = (\partial_\mu \underset{n}{e_\kappa} - \partial_\kappa \underset{n}{e_\mu}) \underset{k}{e^\kappa} \underset{m}{e^\mu},$$

or

$$\underset{knm}{\gamma} - \underset{mnk}{\gamma} = (\partial_\mu \underset{n}{e_\kappa} - \partial_\kappa \underset{n}{e_\mu}) \underset{k}{e^\kappa} \underset{m}{e^\mu}. \tag{7.7-14}$$

Next we multiply (7.7–12) by $\underset{k}{e^\kappa} \underset{m}{e^\mu}$ and perform the summation on repeated indices. We get

$$(\partial_\lambda e_\mu - \partial_\mu e_\lambda) e_\kappa \underset{k}{e^\kappa} \underset{m}{e^\mu} + (\partial_\mu e_\kappa - \partial_\kappa e_\mu) e_\lambda \underset{k}{e^\kappa} \underset{m}{e^\mu} + (\partial_\kappa e_\lambda - \partial_\lambda e_\kappa) e_\mu \underset{k}{e^\lambda} \underset{m}{e^\mu} = 0,$$

where e stands for $\underset{n}{e}$. Supposing $k \neq \mathrm{n}$, $m \neq \mathrm{n}$, the conditions of orthogonality imply

$$e_\kappa \underset{k}{e^\kappa} = 0, \qquad e_\mu \underset{m}{e^\mu} = 0.$$

Hence the above equations reduce to

$$(\partial_\mu e_\kappa - \partial_\kappa e_\mu) e_\lambda \underset{k}{e^\kappa} \underset{m}{e^\mu} = 0.$$

Transvection by e^λ yields

$$(\partial_\mu e_\kappa - \partial_\kappa e_\mu) \underset{k}{e^\kappa} \underset{m}{e^\mu} = 0$$

and by virtue of (7.7–13) we have:

A necessary condition that the congruence $\underset{n}{\boldsymbol{e}}$ *should be normal is expressed by*

$$\boxed{\underset{knm}{\gamma} = \underset{mnk}{\gamma},} \qquad k \neq \mathrm{n},\ m \neq \mathrm{n}. \tag{7.7-15}$$

It can be proved that this condition is also sufficient. The proof requires some results of the general theory of the partial differential equations of the first order.

Suppose now that all congruences of the orthogonal system are normal. Assuming $k \neq l \neq m \neq k$ we find from of (7.7–8) and (7.7–15)

$$\underset{klm}{\gamma} = \underset{mlk}{\gamma} = -\underset{lmk}{\gamma} = -\underset{kml}{\gamma} = \underset{mkl}{\gamma} = \underset{lkm}{\gamma} = -\underset{klm}{\gamma}$$

Hence:

When all congruences of an orthogonal system are normal, then

$$\underset{klm}{\gamma} = 0, \qquad k \neq l \neq m \neq k. \tag{7.7–16}$$

7.7.5 – Canonical congruences

We wish to study an orthogonal system of congruences whose first $n-1$ members are related in a particular fashion to the remaining one. First we recall the expression (7.7–13) for the coefficients of rotation which may be written as

$$\underset{knm}{\gamma} = (\partial_\mu e_\kappa - \Gamma^\nu{}_{\kappa\mu} e_\nu)\underset{k}{e^\kappa}\, \underset{m}{e^\mu} = e_{\kappa\mu}\underset{k}{e^\kappa}\, \underset{m}{e^\mu} \tag{7.7–17}$$

where e_ν stands for $\underset{n}{e_\nu}$. It is easy to verify that the expressions

$$e_{\kappa\mu} = \partial_\mu e_\kappa - \Gamma^\nu{}_{\kappa\mu} e_\nu \tag{7.7–18}$$

are the component of a tensor. In fact, it follows from

$$e_{\kappa'} = e_\kappa \partial_{\kappa'} q^\kappa$$

that

$$\partial_{\mu'} e_{\kappa'} = \partial_\mu e_\kappa \partial_{\mu'} q^\mu \partial_{\kappa'} q^\kappa + e_\kappa \partial_{\kappa'\mu'} q^\kappa.$$

On the other hand we deduce from (7.4–3)

$$\Gamma^{\nu'}{}_{\kappa'\mu'} e_{\nu'} = \Gamma^{\nu'}{}_{\kappa'\mu'} e_\nu \partial_{\nu'} q^\nu e_\nu = \Gamma^\nu{}_{\lambda\mu} e_\nu \partial_{\kappa'} q^\kappa \partial_{\mu'} q^\mu + e_\nu \partial_{\kappa'\mu'} q^\nu$$

and it appears that

$$e_{\kappa'\mu'} = e_{\kappa\mu} \partial_{\kappa'} q^\kappa \partial_{\mu'} q^\mu.$$

Now we form a symmetric tensor with components

$$c_{\kappa\mu} = \tfrac{1}{2}(e_{\kappa\mu} + e_{\mu\kappa}). \tag{7.7–19}$$

It defines an operator C. We observe that with respect to the orthonormal frame defined by the tangent vectors at a given point of the curves passing through this point, an arbitrary vector $\boldsymbol{e}$ may be represented as

$$\boldsymbol{e} = \sum_{m=1}^{n} \underset{m}{e}\, \underset{m}{\boldsymbol{e}}.$$

Since

$$\underset{m}{e} = x_\mu \underset{m}{e}^\mu, \quad m = 1, \ldots, \mathrm{n},$$

we also have

$$e = x_\mu \sum_{m=1}^{\mathrm{n}} \underset{m}{e}^\mu \underset{m}{e} = x_\mu e^\mu.$$

Hence

$$\sum_{m=1}^{\mathrm{n}} \underset{m}{e}^\mu \underset{m}{e} = e^\mu,$$

an equation relating the orthogonal and the general components of the vector e. Taking account of (2.3–16) and (7.7–17) we easily see that the operator C takes the form

$$\begin{aligned}\mathsf{C}e = \mathsf{C}x^\kappa e_\kappa &= \sum_{k=1}^{\mathrm{n}} \underset{k}{e}\, \underset{k}{e}^\kappa c_{\kappa\mu}\, e^\mu = \sum_{k=1}^{\mathrm{n}} \sum_{m=1}^{\mathrm{n}} \underset{k}{e}\, c_{\kappa\mu} \underset{k}{e}^\kappa \underset{m}{e}^\mu \underset{m}{e} \\ &= \tfrac{1}{2}\sum_{k=1}^{\mathrm{n}} \sum_{m=1}^{\mathrm{n}} \underset{k}{e} (\underset{k\mathrm{n}m}{\gamma} + \underset{m\mathrm{n}k}{\gamma}) \underset{m}{e} = \tfrac{1}{2}\sum_{k=1}^{\mathrm{n}} \sum_{h=1}^{\mathrm{n}} \underset{k}{e} (\underset{k\mathrm{n}h}{\gamma} + \underset{h\mathrm{n}k}{\gamma}) \underset{h}{e},\end{aligned} \tag{7.7–20}$$

where e is an arbitrary vector.

We are particularly interested in the vectors orthogonal to $\underset{\mathrm{n}}{e}$. Let P denote the projection operator associated with the space spanned by the vectors $\underset{1}{e}, \ldots, \underset{\mathrm{n}-1}{e}$. If e is a vector included in this space then $\underset{\mathrm{n}}{e} = 0$ and

$$\mathsf{PC}e = \tfrac{1}{2}\sum_{k=1}^{\mathrm{n}} \sum_{h=1}^{\mathrm{n}} \underset{k}{e} (\underset{k\mathrm{n}h}{\gamma} + \underset{h\mathrm{n}k}{\gamma}) \underset{h}{e}, \tag{7.7–21}$$

and it appears that PC is symmetric. Hence at every point of the manifold this operator admits a set of $n-1$ mutually principal directions.

Now we introduce the concept of canonical congruence. A congruence orthogonal to a congruence $\underset{\mathrm{n}}{e}$ is said to be *canonical* with respect to this latter congruence if the directions of its curves are principal directions of the operator PC.

A necessary and sufficient condition for the congruence $\underset{m}{e}$ of the orthogonal system $\underset{1}{e}, \ldots, \underset{\mathrm{n}-1}{e}$ to be canonical with respect to $\underset{\mathrm{n}}{e}$ is that

$$\mathsf{PC}\underset{m}{e} = \underset{m}{\sigma} \underset{m}{e}, \qquad m = 1, \ldots, \mathrm{n}-1. \tag{7.7–22}$$

Replacing e by $\underset{m}{e}$, in (7.7–21) on the right hand side we must take $\underset{h}{e} = 0$ if

$m \neq h$ and $\underset{m}{e} = 1$. Hence, according to (7.7–22)

$$\sum_{k=1}^{n} \underset{k}{e}(\underset{knm}{\gamma} + \underset{mnk}{\gamma}) = \underset{m}{\sigma}\, \underset{m}{e}, \qquad m = 1, \ldots, n-1. \tag{7.7–23}$$

These equations imply

$$\boxed{\underset{knm}{\gamma} + \underset{mnk}{\gamma} = 0,} \qquad m \neq k,\ m, k = 1, \ldots, n-1, \tag{7.7–24}$$

and

$$\boxed{\underset{m}{\sigma} = \underset{mnm}{\gamma},} \qquad m = 1, \ldots, n-1. \tag{7.7–25}$$

Thus we proved:

A necessary and sufficient condition for the first $n-1$ congruences of an orthonormal system to be canonical with respect to the n-th congruence is that the conditions (7.7–24) *be satisfied.*

The operator C applied to $\underset{m}{e}$, $m = 1, \ldots, n-1$ yields

$$\mathsf{C}\underset{m}{e} = \underset{m}{\sigma}\, \underset{m}{e} + \underset{m}{\rho}\, \underset{n}{e} \tag{7.7–26}$$

where, according to (7.7–20)

$$\underset{m}{\rho} = \tfrac{1}{2} \underset{mnn}{\gamma}, \qquad m = 1, \ldots, n-1, \tag{7.7–27}$$

fr $\underset{nnm}{\gamma} = 0$.

The eigenvectors of the operator PC are solutions of the equations

$$\mathsf{PC}e = \sigma e$$

when σ is suitably chosen. Hence

$$\mathsf{C}e = \sigma e + \rho \underset{n}{e}.$$

Expressed in terms of the components with respect to a frame (κ) these equations are equivalent to the system

$$c_{\kappa\mu} e^{\mu} \sigma g_{\kappa\mu} e^{\mu} - \rho \underset{n}{e_{\kappa}} = 0, \qquad \kappa = 1, \ldots n, .$$

Eliminating e^{μ} and ρ from these equations and the condition of orthogonality

$$e^{\mu} \underset{n}{e_{\mu}} = 0$$

we find that the eigenvalues $\underset{h}{\sigma}$ are the roots of the characteristic equation

$$\begin{vmatrix} c_{11}-\sigma g_{11} & \dots & c_{1n}-\sigma g_{1n} & \underset{\mathrm{n}}{e_1} \\ \cdot & \cdot & \cdot & \cdot \\ c_{n1}-\sigma g_{n1} & \dots & c_{nn}-\sigma g_{nn} & \underset{\mathrm{n}}{e_n} \\ \underset{\mathrm{n}}{e_1} & \dots & \underset{\mathrm{n}}{e_n} & 0 \end{vmatrix} = 0. \qquad (7.7\text{–}28)$$

It is easy to verify that this equation is of degree $n-1$, as should be the case.

We wish to conclude this section by stating some direct consequences of the results obtained.

If the n-th congruence is normal, then (7.7–15) holds. Combined with (7.7–24) we get

A necessary and sufficient condition for $n-1$ mutually orthogonal congruences orthogonal to a normal congruence to be canonical with respect to the latter is that

$$\underset{k\mathrm{n}m}{\gamma} = 0 \qquad k \neq m, \quad k \neq \mathrm{n}, \quad m \neq \mathrm{n}. \qquad (7.7\text{–}29)$$

Taking account of (7.7–16) we have in addition:

When a manifold admits an orthogonal system of n normal congruences, then any of these is canonical with respect to each other congruence of the system.

CHAPTER 8

HYPERSURFACES

The investigation of the properties of a hypersurface is much facilitated by the fact that a manifold of this kind possesses a field of normals such that through every (regular) point passes just one line orthogonal to the manifold. This field enables us to penetrate directly into the theory of the curvature.

It is clear that properties derived by means of the field of normals depend on the shape of the hypersurface. It is noteworthy, therefore, that there are theorems about curvature based on the properties of the field of normals which turn out to be independent of the shape of the hypersurface, provided it is isometrically deformed. These so-called intrinsic properties are very important, for they lend themselves to generalization to abstract metric spaces, being point sets endowed with a metric, but not necessarily embedded in a linear space.

The subject-matter dealt with in this chapter is dominated by the theory of the system of normals. A striking feature is the fact that even the most elementary parts of the theory of vectors are already sufficient to exhibit interesting geometric facts. Afterwards we shall be able to generalize a famous result due to Gauss, which provides us with a metric invariant and brings us in contact with an important tensor discovered by Riemann.

8.1 – The field of normals

8.1.1 – EXISTENCE

In a metric space of dimension $n+1$ we consider an n-dimensional manifold, a *hypersurface*, characterized by the vector function

$$\boldsymbol{x}(q^\kappa), \qquad \kappa = 1, \ldots, n. \tag{8.1–1}$$

As usual, we suppose that n is at least two. It is our aim to establish the existence of a differentiable field

$$\boldsymbol{n}(q^\kappa), \qquad \kappa = 1, \ldots, n, \tag{8.1–2}$$

such that

$$\boldsymbol{n}\partial_\kappa \boldsymbol{x} = 0, \qquad \kappa = 1, \ldots, n, \tag{8.1–3}$$

and

$$\boldsymbol{n}\boldsymbol{n} = 1. \qquad (8.1\text{–}4)$$

To this end we introduce in the space an orthonormal frame $\underset{h}{\boldsymbol{e}}$, $h = 1, \ldots, n+1$. Referring to this frame we may write

$$\boldsymbol{n} = \sum_{h=1}^{n+1} \underset{h}{\boldsymbol{e}} \underset{h}{n},$$

and

$$\partial_\kappa \boldsymbol{x} = \sum_{h=1}^{n+1} \underset{h}{\boldsymbol{e}} \underset{h}{x_\kappa} \qquad \kappa = 1, \ldots, n.$$

Then (8.1–3) assumes the form

$$\sum_{h=1}^{n+1} \underset{h}{n} \underset{h}{x_\kappa} = 0,$$

a system of n homogeneous equations in $n+1$ unknowns. By hypothesis the rank of the matrix

$$\begin{bmatrix} \underset{1}{x_1} & \underset{n+1}{x_1} \\ \underset{1}{x_n} & \underset{n+1}{x_n} \end{bmatrix} \qquad (8.1\text{–}5)$$

equals n. The unknowns $\underset{h}{n}$ are proportional to the determinants of order n taken from this matrix and provided with an appropriate sign. Since not all these determinants are zero we can take the factor of proportionality such that (8.1–4) has also been satisfied. Moreover, it turns out that the vector function (8.1–2) is of class C^u provided (8.1–1) is of class C^{u+1}. In the subsequent considerations we suppose this class to be sufficiently high in order to assure the validity of the results.

A vector system satisfying (8.1–3) and (8.1–4) is called a *field of normals*. It should be noticed, that the field of normals is determined except for sign. In most cases, however, the ambiguity in sign does not cause any trouble.

8.1.2 – The hypersphere

A simple example of a hypersurface is provided by the *hypersphere*, being the set of points at a distant ρ from a given point. The (positive) number ρ is the radius of the hypersphere. When the *centre* is characterized by the vector $\boldsymbol{a}$, then the hypersphere is the set of points whose coordinate vector $\boldsymbol{x}$ satisfies the equation

$$(\boldsymbol{x}-\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a}) = \rho^2. \qquad (8.1\text{–}6)$$

Without loss of generality we may suppose that $\boldsymbol{a} = \boldsymbol{o}$, i.e., the origin of the coordinate system in the space is the centre of the hypersphere.

It is not possible to represent the whole hypersphere by a function $\boldsymbol{x}(q^\kappa)$ satisfying the conditions listed in section 5.3.1. But, since we are only interested in local properties, this fact does not disturb us, for we may confine ourselves to a part of the hypersphere. Consider a hyperplane through the centre of the hypersphere given by the equation

$$\boldsymbol{xp} = 0, \tag{8.1–7}$$

where $\boldsymbol{p}$ denotes a fixed vector in space. The points of the hypersphere with $\boldsymbol{xp} > 0$ are on one side of this hyperplane and constitute a set, which is in one-to-one correspondence with a region in the hyperplane, viz., the region obtained by projecting orthogonally the points of the hypersphere on the hyperplane. Now we have

$$\boldsymbol{x} = \boldsymbol{y}+\boldsymbol{p}z, \tag{8.1–8}$$

where $\boldsymbol{y}$ is the projection of $\boldsymbol{x}$ on the hyperplane. Assuming that $\boldsymbol{p}$ is a unit vector we may infer from

$$\boldsymbol{xx} = \rho^2 \tag{8.1–9}$$

that

$$z = \sqrt{\rho^2-\boldsymbol{yy}}, \qquad \boldsymbol{yy} < \rho^2.$$

The positive sign of the square root implies $\boldsymbol{xp} = z > 0$. On the other hand we may suppose that $\boldsymbol{y}$ is a function of certain parameters $q^1, \ldots, q^n$ in the hyperplane of sufficiently high class, for instance of the rectangular coordinates with respect to an orthonormal frame. As a consequence z and $\boldsymbol{x}$ can also be expressed in terms of these parameters and turn out to be functions of the same class.

On differentiating (8.1–7) we obtain

$$\partial_\kappa \boldsymbol{x} = \partial_\kappa \boldsymbol{y}+\boldsymbol{p}\partial_\kappa z.$$

A linear relation

$$\partial_\kappa \boldsymbol{x}\, \xi^\kappa = \boldsymbol{o}$$

implies

$$\partial_\kappa \boldsymbol{y}\, \xi^\kappa+\boldsymbol{p}\, \partial_\kappa z \xi^\kappa = \boldsymbol{o}.$$

Since $\boldsymbol{py} = 0$ we also have $\boldsymbol{p}\partial_\kappa \boldsymbol{y} = 0$ and the above equation implies $\xi^\kappa\, \partial_\kappa z = 0$. Hence

$$\partial_\kappa \boldsymbol{y} \xi^\kappa = \boldsymbol{o}.$$

But we may suppose that the system of vectors $\partial_\kappa \boldsymbol{y}$ is linearly independent.

Hence $\xi^\kappa = 0$, $\kappa = 1, \ldots, n$ and thus we proved that a semihypersphere is a hypersurface.

On differentiating (8.1–9) we get

$$\boldsymbol{x}\, \partial_\kappa \boldsymbol{x} = 0, \qquad \kappa = 1, \ldots, n. \tag{8.1–10}$$

Hence the vector $\boldsymbol{x}$ is a solution of (8.1–3). It is customary to norm this vector in such a way that

$$\boldsymbol{n} = -\boldsymbol{x}\,(1/\rho). \tag{8.1–11}$$

that is to say, the normal points to the interior of the hypersphere.

It follows at once from the last theorem of section 7.2.2 that a great circle on a hypersphere — being the intersection of the hypersphere and a plane through the centre — is a geodesic line. Through a given point there passes just one great circle having there a prescribed direction. Hence, according to the last theorem of section 7.5.4 the great circles are the only geodesics on a hypersphere.

8.1.3 – Curvature and torsion

It is natural to investigate the rate of change of a normal vector along a curve $q^\kappa(s)$, $\kappa = 1, \ldots, n$ on the hypersurface, the curve being related to its arc length as parameter. On differentiating (8.1–4) we find

$$\boldsymbol{n}\, \mathrm{d}_s \boldsymbol{n} = 0. \tag{8.1–12}$$

Hence *the vector* $\mathrm{d}_s \boldsymbol{n}$ *is a tangential vector.*

Let $\boldsymbol{u}$ denote the tangent vector

$$\boldsymbol{u} = \mathrm{d}_s \boldsymbol{x} \tag{8.1–13}$$

of the given curve $\boldsymbol{x}(s)$ on the hypersurface. We may write

$$\mathrm{d}_s \boldsymbol{n} = -\boldsymbol{u}\kappa + \boldsymbol{v}\tau \tag{8.1–14}$$

where $\boldsymbol{v}$ is a unit vector orthogonal to $\boldsymbol{u}$ (and also contained in the tangent space at the point under consideration). We agree to take $\boldsymbol{v}$ such that τ is always non-negative.

The number κ, which does not have a fixed sign, is called *the curvature of the hypersurface in the direction* $\boldsymbol{u}$. It is not unambiguously defined, for it depends on the choice of the field of normals. Strictly speaking, curvature is not a property of the hypersurface, but of the combination of a hypersurface and one of its fields of normals.

The number τ is called the *torsion of the hypersurface in the direction* $\boldsymbol{u}$. According to the definition it is unambiguously determined.

It may happen that $\kappa = 0$. Then the direction is called an *asymptotic direction.* A curve having at each point an asymptotic direction is called an *asymptotic curve.*

Not every hypersurface possesses asymptotic curves. In the case of a hypersphere, for instance, we always have

$$\mathrm{d}_s \boldsymbol{n} = - \mathrm{d}_s \boldsymbol{x}(1/\rho) \tag{8.1–15}$$

as follows from (8.1–11). Hence $\kappa = 1/\rho \neq 0$. Thus we see that on a hypersphere there are no asymptotic curves. On the other hand every curve in a hyperplane is an asymptotic curve. In fact, (8.1–7) implies $\boldsymbol{p}\partial_\kappa \boldsymbol{x} = 0$, and we may, therefore, take $\boldsymbol{n} = \boldsymbol{p}$. But since $\boldsymbol{p}$ is constant we have $\mathrm{d}_s\boldsymbol{n} = \boldsymbol{o}$.

From now on asymptotic curves will not concern us. We wish to turn our attention to curves having at each point a direction in which the torsion vanishes. Such a direction is referred to as a *principal direction of curvature,* briefly a *principal direction.* A curve having at each point a principal direction is called a *principal curve* or a *line of curvature.* The existence of principal directions will be established in section 8.2.3. The curvatures in the principal directions play an important part in the theory. They are called the *principal curvatures.*

8.1.4 – Lines of curvature

A line of curvature is characterized by the property that the rate of change of the normals along this curve is in the direction of the curve. Otherwise stated:

A curve on a hypersurface is a line of curvature if and only if the rate of change of the normals along the curve satisfies the condition

$$\boxed{\mathrm{d}_s \boldsymbol{n} + \mathrm{d}_s \boldsymbol{x}\, \kappa = 0.} \tag{8.1–16}$$

In this equation, the *equation of Olinde-Rodrigues,* we may replace the arc length s by an arbitrary parameter q.

This equation admits a very interesting geometric interpretation. To this end we need the concept of *developable surface,* being a surface described by the tangent lines of a curve in space, a cone, or a cylinder. In these latter cases the lines rest on a curve and pass either through a fixed point or are parallel to a fixed line. This concept enables us to characterize the lines of curvature in the following way:

A necessary and sufficient condition for a curve on a hypersurface to be a line of curvature is that the lines through the points of the curve orthogonal to the hypersurface constitute a developable surface.

Thus, for instance, every curve on a hypersphere is a line of curvature, for the normal lines all pass through the centre of the hypersphere. Also every curve in a hyperplane is a line of curvature, for the lines through the points of this curve are the generators of a cylinder.

In order to prove the theorem we consider the vector function

$$\boldsymbol{y}(q) = \boldsymbol{x}(q)+\boldsymbol{n}(q)\beta(q) \tag{8.1–17}$$

depending on the parameter q along the curve. By β we denote a function of this parameter which is to be determined appropriately.

First we suppose that the given curve is a line of curvature. On account of (8.1–16) we have

$$\begin{aligned}\mathrm{d}_q\boldsymbol{y} &= \mathrm{d}_q\boldsymbol{x}+\mathrm{d}_q\boldsymbol{n}\,\beta+\boldsymbol{n}\,\mathrm{d}_q\beta\\ &= \mathrm{d}_q\boldsymbol{x}(1-\kappa\beta)+\boldsymbol{n}\,\mathrm{d}_q\beta,\end{aligned}$$

for, as we already pointed out, the parameter s occurring in (8.1–16) may be replaced by any other parameter.

1) Let $\kappa \neq 0$. Then we take $\beta = 1/\kappa$ and the above equation reduces to

$$\mathrm{d}_q\boldsymbol{y} = \boldsymbol{n}\,\mathrm{d}_q\beta. \tag{8.1–18}$$

If κ is not constant then $\boldsymbol{y}$ characterizes a curve, whose tangent lines according to (8.1–18), are parallel to $\boldsymbol{n}$. If, however, κ = constant, then $\mathrm{d}_q\beta = 0$ and $\boldsymbol{y}$ is constant. This means that all normals to the hypersurface at the points of the given curve pass through a fixed point and are, therefore, the generators of a cone.

2) Let $\kappa = 0$ along the curve. Then (8.1–16) expresses the fact that $\boldsymbol{n}$ is constant along the curve, and it follows that the normal lines are the generators of a cylinder.

To establish the sufficiency of the condition we must distinguish three cases:

1) The normal lines are tangents of a space curve $\boldsymbol{y}(q)$. The lines connecting the corresponding points of the curves $\boldsymbol{y}(q)$ and $\boldsymbol{x}(q)$, being the points with the same parameter, are parallel to the normal. Hence (8.1–17) holds and on differentiating we get

$$\mathrm{d}_q\boldsymbol{y} = \mathrm{d}_q\boldsymbol{x}+\boldsymbol{n}\,\mathrm{d}_q\beta+\mathrm{d}_q\boldsymbol{n}\,\beta. \tag{8.1–19}$$

This vector is a tangent vector of the curve $\boldsymbol{y}(q)$ and since it depends linearly on $\boldsymbol{n}$ we can find a function λ such that

$$\mathrm{d}_q \boldsymbol{x}+\mathrm{d}_q \boldsymbol{n}\,\beta = \boldsymbol{n}\lambda. \tag{8.1–20}$$

By multiplying both members by $\boldsymbol{n}$ we see that $\lambda = 0$ and it follows that (8.1–16) holds with $\kappa = 1/\beta$, provided $\beta \neq 0$. But $\beta = 0$ is impossible, for $\mathrm{d}_q \boldsymbol{x} \neq \boldsymbol{o}$.

2) If the normal lines along the given curve constitute a cone, then again (8.1–17) holds, where $\boldsymbol{y}$ is now a constant. In (8.1–19) we have $\mathrm{d}_q \boldsymbol{y} = \boldsymbol{o}$ and on multiplying this equation by $\boldsymbol{n}$ we see that $\mathrm{d}_q \beta = 0$, i.e., β is constant. As a consequence (8.1–20) is valid with $\lambda = 0$ and this proves the theorem for this case.

3) If the normal lines constitute a cylinder then $\mathrm{d}_q \boldsymbol{n} = \boldsymbol{o}$ and (8.1–16) holds for $\kappa = 0$.

Another direct consequence of the equation of Olinde-Rodrigues is the following *theorem of Joachimsthal*:

If a curve C, common to two hypersurfaces, is a line of curvature of either hypersurface then the hypersurfaces meet at a constant angle along C.

Let $\boldsymbol{x}(q^\kappa)$ and $\overset{\bullet}{\boldsymbol{x}}(q^\kappa)$ characterize the hypersurfaces. Along C we have $\boldsymbol{x} = \overset{\bullet}{\boldsymbol{x}}$. We denote the principal curvature along C on the first hypersurface by κ and on the second hypersurface by $\overset{\bullet}{\kappa}$. If $\boldsymbol{n}$ and $\overset{\bullet}{\boldsymbol{n}}$ are the normals of the two hypersurfaces at a point of C, the angle θ mentioned in the theorem is defined by

$$\cos\theta = \boldsymbol{n}\overset{\bullet}{\boldsymbol{n}}.$$

On differentiating we get

$$\begin{aligned}\mathrm{d}_q \cos\theta &= \mathrm{d}_q(\boldsymbol{n}\overset{\bullet}{\boldsymbol{n}}) = \boldsymbol{n}\,\mathrm{d}_q\overset{\bullet}{\boldsymbol{n}}+\overset{\bullet}{\boldsymbol{n}}\,\mathrm{d}_q\boldsymbol{n}\\ &= -\overset{\bullet}{\kappa}\boldsymbol{n}\,\mathrm{d}_q\boldsymbol{x}-\kappa\overset{\bullet}{\boldsymbol{n}}\,\mathrm{d}_q\boldsymbol{x}\\ &= -\overset{\bullet}{\kappa}\boldsymbol{n}\,\mathrm{d}_q\boldsymbol{x}-\kappa\overset{\bullet}{\boldsymbol{n}}\,\mathrm{d}_q\overset{\bullet}{\boldsymbol{x}} = 0.\end{aligned}$$

This proves the assertion.

8.1.5 – Equidistant hypersurfaces

If on the normal lines of a hypersurface a segment of constant length is laid off pointing to the same side and having a point of the hypersurface as initial point, then the locus of the end points is, in general, again a hypersurface, as we shall see presently. This locus is more precisely described by the function

$$\hat{\boldsymbol{x}} = \boldsymbol{x}+\boldsymbol{n}\beta, \tag{8.1–21}$$

where β is a constant. It is a hypersurface if and only if the vectors $\partial_\kappa \hat{\boldsymbol{x}}$ are linearly independent. Suppose they are linearly dependent. Then we have a relation

$$\partial_\kappa \boldsymbol{x}\, u^\kappa + \partial_\kappa \boldsymbol{n}\, u^\kappa \beta = \boldsymbol{o} \tag{8.1–22}$$

and without loss of generality we may suppose that the u^κ are the components $\mathrm{d}_s q^\kappa$ of a tangent vector $\boldsymbol{u}$ of a curve $q^\kappa(s)$. Hence we may write (8.1–22) in the form

$$\boldsymbol{u} + \mathrm{d}_s \boldsymbol{n}\, \beta = \boldsymbol{o}. \tag{8.1–23}$$

Comparing this with (8.1–14) we see that the torsion in the direction $\boldsymbol{u}$ is zero and, as a consequence, β is the reciprocal of a principal curvature. Thus we see:

The function (8.1–21) *represents a hypersurface if and only if* $1/\beta$ *is not a principal curvature.*

The hypersurfaces $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ are called *equidistant or parallel.*

Two points on equidistant hypersurfaces lying on the same normal line are considered as corresponding points. A direct consequence of the first theorem of the previous section states:

On two equidistant hypersurfaces the lines of curvature correspond.

It is not difficult to evaluate the principal curvatures of the hypersurface $\hat{\boldsymbol{x}}$, when those of $\boldsymbol{x}$ are known. Let $\mathrm{d}_q \boldsymbol{x}$ denote a vector pointing in a principal direction. From (8.1–21) and the equation of Olinde-Rodrigues it follows that

$$\mathrm{d}_q \hat{\boldsymbol{x}} = \mathrm{d}_q \boldsymbol{n}(1 - \kappa^{-1} + \beta)$$

provided $\kappa \neq 0$. The principal curvature at the corresponding point of the equidistant hypersurface (8.1–21) is, therefore, determined by

$$\hat{\kappa}^{-1} = \kappa^{-1} - \beta$$

or

$$\hat{\kappa} = \frac{\kappa}{1 - \beta\kappa}. \tag{8.1–24}$$

The denominator is not zero, for $\beta\kappa \neq 1$ as we already pointed out. The formula (8.1–24) remains valid for $\kappa = 0$; then also $\hat{\kappa} = 0$.

8.1.6 – Evolutes

It is also natural to consider manifolds $\check{\boldsymbol{x}}$ of the type (8.1–21), where β designates a certain function of the parameters q^κ, not necessarily being constant. Then

$$\partial_\lambda \check{\boldsymbol{x}} = \partial_\lambda \boldsymbol{x} + \partial_\lambda \boldsymbol{n}\, \beta + \boldsymbol{n} \partial_\lambda \beta. \tag{8.1–25}$$

When the vectors $\partial_\lambda \check{\boldsymbol{x}}$, $\lambda = 1, \ldots, n$, are linearly dependent at a given point, we can find a unit vector $\boldsymbol{u}$ such that

$$\boldsymbol{o} = \mathrm{d}_s \check{\boldsymbol{x}} = \boldsymbol{u} + \mathrm{d}_s \boldsymbol{n}\, \beta + \boldsymbol{n}\, \mathrm{d}_s \beta,$$

where $\boldsymbol{u} = \partial_\lambda \boldsymbol{x} u^\lambda$, $\mathrm{d}_s \check{\boldsymbol{x}} = \partial_\lambda \check{\boldsymbol{x}} u^\lambda$, $\mathrm{d}_s \boldsymbol{n} = \partial_\lambda \boldsymbol{n} u^\lambda$ and $\mathrm{d}_s \beta = \partial_\lambda \beta u^\lambda$. It follows by multiplication by $\boldsymbol{n}$ that $\mathrm{d}_s \beta = 0$, i.e., β is stationary in the direction $\boldsymbol{u}$. Thus we may state: *Under the assumption that* $\mathrm{d}_s \beta$ *vanishes nowhere, the differentiation being performed along an arbitrary curve on the hypersurface, the vector function* (8.1–21) *defines another hypersurface referred to the parameters* q^κ *of the original hypersurface.*

We wish to discuss in more detail a particular case, by taking for β the reciprocal of a principal curvature. First we wish to complete slightly the definitions of section 8.1.4. A function κ is called a *principal curvature* when at every point of the hypersurface a direction $\mathrm{d}_s \boldsymbol{x}$ can be found, such that (8.1–16) holds, this direction being a principal direction corresponding to κ. If at every point of a curve the direction is a principal direction, then the curve is a line of curvature corresponding to κ. In section 8.1.4 we proved that the normal lines along a line of curvature constitute a developable surface. Excluding the case that κ presents stationary values, we may assert that the developable surface possesses an edge of regression and this will be called an edge of regression belonging to the line of curvature under consideration.

Next we take $\beta = \kappa^{-1}$. In the light of the results obtained before, all edges of regression belonging to lines of curvature corresponding to κ are situated on a hypersurface

$$\check{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{n}\rho \tag{8.1–26}$$

with $\rho = \kappa^{-1}$. The hypersurface (8.1–26) is called an *evolute* of the given hypersurface associated with the curvature κ. Now we may prove the following interesting theorem:

The edges of regression belonging to lines of curvature corresponding to the principal curvature κ *are geodesics on the evolute associated with* κ.

According to section 7.2.2 we have only to verify that the principal normal of an edge of regression mentioned in the theorem is orthogonal to the evolute. The tangent vector is the normal $\boldsymbol{n}$, the principal normal is, therefore, $\mathrm{d}_s \boldsymbol{n}$, where the differentiation is performed in a principal direction $\boldsymbol{u}$ corresponding to κ. In view of the equation of Olinde-Rodrigues (8.6–1) we have

$$\mathrm{d}_s \boldsymbol{n} = -\boldsymbol{u}\kappa.$$

Hence

$$\partial_\lambda \check{\boldsymbol{x}}\, \mathrm{d}_s \boldsymbol{n} = -\boldsymbol{u}\partial_\lambda \boldsymbol{x}\, \kappa - \boldsymbol{u}\partial_\lambda \boldsymbol{n}.$$

From (8.1–3) we deduce

$$\partial_\lambda \boldsymbol{n} \partial_\mu \boldsymbol{x} = -\boldsymbol{n}\partial_{\lambda\mu}\boldsymbol{x} = \partial_\mu \boldsymbol{n}\partial_\lambda \boldsymbol{x},$$

whence

$$\boldsymbol{u}\partial_\lambda \boldsymbol{n} = \partial_\lambda \boldsymbol{n}\partial_\mu \boldsymbol{x}\, u^\mu = \partial_\lambda \boldsymbol{x}\partial_\mu \boldsymbol{n}\, u^\mu = \partial_\lambda \boldsymbol{x}\, \mathrm{d}_s \boldsymbol{n}.$$

Thus we find

$$\partial_\lambda \check{\boldsymbol{x}}\, \mathrm{d}_s \boldsymbol{n} = -\partial_\lambda \boldsymbol{x}(\boldsymbol{u}\kappa + \mathrm{d}_s \boldsymbol{n}) = 0,$$

referring again to the equation of Olinde-Rodrigues. This completes the proof of the theorem.

8.2 – The second fundamental tensor

8.2.1 – The curvature

On multiplying both members of (8.1–14) by the unit vector $\boldsymbol{u}$ we find the following expression for κ:

$$\boxed{\kappa = -\boldsymbol{u}\, \mathrm{d}_s \boldsymbol{n} = h_{\lambda\mu} u^\lambda u^\mu} \tag{8.2–1}$$

with

$$\boxed{h_{\lambda\mu} = -\partial_\lambda \boldsymbol{x}\partial_\mu \boldsymbol{n}.} \tag{8.2–2}$$

Since $\boldsymbol{x}$ and $\boldsymbol{n}$ are vectorial invariants it follows at once, that the $h_{\lambda\mu}$ are the coefficients of a tensor, called the *second fundamental tensor*. In this connection the metric tensor on a hypersurface is called the *first fundamental tensor*. From (8.1–3) follows an alternative representation of this tensor, for

$$\boxed{\boldsymbol{n}\, \partial_{\lambda\mu} \boldsymbol{x} = -\partial_\lambda \boldsymbol{x}\, \partial_\mu \boldsymbol{n},}$$

i.e.,

$$\boxed{h_{\lambda\mu} = \boldsymbol{n}\partial_{\lambda\mu}\boldsymbol{x}.} \tag{8.2–3}$$

Hence:

The second fundamental tensor is symmetric.

There is an alternative approach to the concept of curvature on a hypersurface. Consider a curve $q^\kappa(s)$ referred to its arc length. As usual we denote the coordinate vector defining this curve by $\boldsymbol{x}(s)$. Differentiation with respect to s yields

$$\mathrm{d}_s \boldsymbol{x} = \partial_\lambda \boldsymbol{x}\, \mathrm{d}_s q^\lambda = \partial_\lambda \boldsymbol{x}\, u^\lambda$$

and differentiating again

$$\mathrm{d}_s{}^2 \boldsymbol{x} = \partial_{\lambda\mu} \boldsymbol{x}\, u^\lambda u^\mu + \partial_\kappa \boldsymbol{x}\, \mathrm{d}_s{}^2 q^\kappa. \tag{8.2–4}$$

The vector $\mathrm{d}_s{}^2 \boldsymbol{x}$ is the (spatial) curvature vector of the curve under consideration (see section 6.1.2). Its component in the direction of $\boldsymbol{n}$ is

$$\boldsymbol{n}\, \mathrm{d}_s{}^2 \boldsymbol{x} = \boldsymbol{n}\, \partial_{\lambda\mu} \boldsymbol{x}\, u^\lambda u^\mu = h_{\lambda\mu} u^\lambda u^\mu = \kappa, \tag{8.2–5}$$

that is, the curvature of the hypersurface in the direction of the curve at the point under consideration. Hence this normal component depends only on this direction.

Denoting by $\tilde{\kappa}$ the curvature of the curve, that is the length of the curvature vector $\mathrm{d}_s{}^2 \boldsymbol{x}$, we find from (8.2–5)

$$\boxed{\kappa = \tilde{\kappa} \cos \theta,} \tag{8.2–6}$$

where θ denotes the angle between $\boldsymbol{n}$ and $\mathrm{d}_s{}^2 \boldsymbol{x}$. This result is known as *Meusnier's theorem.*

Thus the curvature of a hypersurface can be interpreted as the curvature of the plane section through the normal at the given point having the given direction there.

8.2.2 – The equations of Weingarten

On differentiating (8.1–4) partially we get

$$\boldsymbol{n}\, \partial_\lambda \boldsymbol{n} = 0. \tag{8.2–7}$$

Hence the vectors $\partial_\lambda \boldsymbol{n}$ are tangential vectors and we may write

$$\partial_\lambda \boldsymbol{n} = \boldsymbol{x}^\kappa \eta_{\kappa\lambda}.$$

By multiplying both members by $\boldsymbol{x}_\mu$ we arrive at

$$-h_{\lambda\mu} = \partial_\lambda \boldsymbol{n}\, \partial_\mu \boldsymbol{x} = \delta^\kappa_\mu \eta_{\kappa\lambda} = \eta_{\mu\lambda}.$$

Thus we have proved *Weingarten's equations*:

$$\boxed{\partial_\lambda \boldsymbol{n} = -\boldsymbol{x}^\kappa h_{\kappa\lambda}.} \tag{8.2–8}$$

They enable us to interpret the linear transformation defined by the second fundamental tensor. In fact, let $\boldsymbol{u}$ denote a unit vector with components $u^\lambda = \mathrm{d}_s q^\lambda$, the differentiation being performed along a curve having the direction $\boldsymbol{u}$ at the given point. Then

$$\mathrm{d}_s \boldsymbol{n} = -\boldsymbol{x}^\kappa h_{\kappa\lambda} u^\lambda$$

denotes a vector with components $-h_{\kappa\lambda} u^\lambda$ representing the rate of change of the field of normal vectors in the direction $\boldsymbol{u} = \boldsymbol{x}_\kappa u^\kappa$. This means that the vector $\mathrm{d}_s \boldsymbol{x} = \boldsymbol{u}$ is transformed into the vector $-\mathrm{d}_s \boldsymbol{n}$ when we apply the operator H with components $h_{\lambda\mu}$.

8.2.3 – The principal curvatures as eigenvalues

The last remark of the preceding section, combined with the equation of Olinde-Rodrigues, tells us that the principal curvatures on a hypersurface at a point are the eigenvalues of the second fundamental tensor evaluated at this point. Hence they are the roots of the characteristic equation

$$\det\,[h_{\lambda\mu} - \kappa g_{\lambda\mu}] = 0. \tag{8.2–9}$$

It also follows that the principal directions always exist and that at every point we can find an orthonormal system of principal directions.

Another consequence is the following. Let $\underset{h}{\theta}$ denote the angles of an arbitrary direction with the principal directions, where h runs from 1 to n. Then, according to (3.2–3), the curvature κ in this direction may be expressed in terms of the principal curvatures $\underset{h}{\kappa}$, $h = 1, \ldots, \mathrm{n}$, by means of *Euler's formula*

$$\boxed{\kappa = \sum_{h=1}^{\mathrm{n}} \underset{h}{\kappa} \cos^2 \underset{h}{\theta}.} \tag{8.2–10}$$

8.2.4 – The third fundamental tensor

The norm of the vector $\mathrm{d}_s \boldsymbol{n}$ is evidently

$$\mathrm{d}_s \boldsymbol{n}\, \mathrm{d}_s \boldsymbol{n} = \partial_\lambda \boldsymbol{n}\, \partial_\mu \boldsymbol{n}\, u^\lambda u^\mu \tag{8.2–11}$$

with $u^\kappa = \mathrm{d}_s q^\kappa$. The expressions

$$\boxed{k_{\lambda\mu} = \partial_\lambda \boldsymbol{n}\, \partial_\mu \boldsymbol{n}} \tag{8.2–12}$$

are evidently the covariant components of a symmetric tensor, the *third fundamental tensor*. It gives information about the torsion in the direction $\boldsymbol{u}$, for by virtue of (8.1–14) we have

$$\boxed{k_{\lambda\mu} u^\lambda u^\mu = \kappa^2 + \tau^2,} \tag{8.2–13}$$

this being the norm of the vector $\mathrm{d}_s \boldsymbol{n} = \partial_\lambda \boldsymbol{n}\, u^\lambda$.

It is possible to express the components of the third fundamental tensor in terms of the components of the second fundamental tensor. In fact, from Weingarten's equations we deduce at once

$$k_{\lambda\mu} = \partial_\lambda \boldsymbol{n}\, \partial_\mu \boldsymbol{n} = \boldsymbol{x}^\alpha h_{\alpha\lambda} \boldsymbol{x}^\beta h_{\beta\mu},$$

or

$$\boxed{k_{\lambda\mu} = g^{\alpha\beta} h_{\alpha\lambda} h_{\beta\mu}.} \tag{8.2–14}$$

The third fundamental tensor is, therefore, the iteration of the second fundamental tensor (compare section 3.3.2). Hence:

The squares of the principal curvatures are the eigenvalues of the third fundamental tensor.

In addition to (8.2–10) on applying (3.2–3) we have

$$\boxed{\kappa^2 + \tau^2 = \sum_{h=1}^{n} \underset{h}{\kappa}^2 \cos^2 \underset{h}{\theta}.} \tag{8.2–15}$$

Next we wish to derive an expression for the metric tensor of a hypersurface (8.1–21) equidistant to a hypersurface $\boldsymbol{x}(q^\kappa)$. Since

$$\partial_\kappa \hat{\boldsymbol{x}} = \partial_\kappa \boldsymbol{x} + \partial_\kappa \boldsymbol{n}\, \beta$$

we infer from (8.2–2) and (8.2–12) that

$$\boxed{\hat{g}_{\lambda\mu} = g_{\lambda\mu} - 2\beta h_{\lambda\mu} + \beta^2 k_{\lambda\mu}.} \tag{8.2–16}$$

With reference to section 5.3.4 we have, in view of the above results:

The principal distortions of the correspondence between the two hypersurfaces $\hat{\boldsymbol{x}}$ *and* $\boldsymbol{x}$ *are*

$$\underset{h}{\omega} = (1 - \beta \underset{h}{\kappa})^2, \qquad h = 1, \ldots, \text{n}. \tag{8.2–17}$$

Agreeing to take $\hat{\boldsymbol{n}} = \boldsymbol{n}$ as the normal for the hypersurface $\hat{\boldsymbol{x}}$ we readily find

$$\hat{h}_{\lambda\mu} = -\partial_\lambda \hat{\boldsymbol{n}}\, \partial_\mu \hat{\boldsymbol{x}} = -(\partial_\lambda \boldsymbol{n} \partial_\mu \boldsymbol{x} + \partial_\lambda \boldsymbol{n}\, \partial_\mu \boldsymbol{n} \beta)$$

i.e.,

$$\boxed{\hat{h}_{\lambda\mu} = h_{\lambda\mu} - \beta k_{\lambda\mu}.} \tag{8.2–18}$$

It is an easy matter to obtain from (8.2–18) expressions for the principal curvatures on the hypersurface $\hat{\boldsymbol{x}}$ in terms of those on $\boldsymbol{x}$. First we observe that a principal direction $\boldsymbol{u} = \boldsymbol{x}_\kappa u^\kappa$ gives rise to an eigenvector $\hat{\boldsymbol{u}} = \hat{\boldsymbol{x}}_\kappa u^\kappa$

on $\hat{x}$ of the operator $\hat{h}_{\lambda\mu}$. The norm of $\hat{u}$ is

$$\hat{u}\hat{u} = \hat{g}_{\lambda\mu} u^\lambda u^\mu = 1-2\beta\kappa+\beta^2\kappa^2 = (1-\beta\kappa)^2.$$

Hence

$$\hat{\kappa} = \frac{\hat{h}_{\lambda\mu} u^\lambda u^\mu}{\hat{g}_{\lambda\mu} u^\lambda u^\mu} = \frac{\kappa-\beta\kappa^2}{(1-\beta\kappa)^2} = \frac{\kappa}{1-\beta\kappa},$$

in accordance with (8.1–24).

8.2.5 – Umbilical points

A point where the principal directions are wholly indeterminate, i.e., where all principal curvatures have the same value, is called an *umbilical point*. It is easy to see that all points of a hyperplane or of a hypersphere are umbilical points. Now we shall establish the converse of this statement, viz.:

A hyperplane and a hypersphere are the only hypersurfaces whose points are all umbilical points.

By the equation of Olinde-Rodrigues the fact that all points are umbilical points can be expressed in the equations

$$\partial_\lambda \boldsymbol{n} + \partial_\lambda \boldsymbol{x}\kappa = 0 \qquad (8.2\text{–}19)$$

where κ is a scalar invariant. On differentiating we get

$$\partial_{\lambda\mu} \boldsymbol{n} + \partial_{\lambda\mu} \boldsymbol{x}\kappa + \partial_\lambda \boldsymbol{x}\, \partial_\mu \kappa = \boldsymbol{o} \qquad (8.2\text{–}20)$$

whence, since the first two terms are symmetric with respect to λ and μ:

$$\partial_\lambda \boldsymbol{x}\, \partial_\mu \kappa - \partial_\mu \boldsymbol{x}\, \partial_\lambda \kappa = \boldsymbol{o}. \qquad (8.2\text{–}21)$$

Transvection by $\boldsymbol{x}^\lambda$ yields

$$n\partial_\mu \kappa - \delta^\lambda_\mu \partial_\lambda \kappa = (n-1)\partial_\mu \kappa = 0, \qquad \mu = 1, \ldots, n,$$

and this means that κ is constant throughout the hypersurface. As a consequence (8.2–19) is equivalent to

$$\boldsymbol{n} + \boldsymbol{x}\kappa = \boldsymbol{p}$$

where $\boldsymbol{p}$ is a constant vector. If $\kappa = 0$ then $\boldsymbol{n} = \boldsymbol{p}$ and $0 = \boldsymbol{xn} = \boldsymbol{xp}$. This characterizes a hyperplane. If $\kappa \neq 0$ then

$$1 = \boldsymbol{nn} = (\boldsymbol{x}\kappa - \boldsymbol{p})(\boldsymbol{x}\kappa - \boldsymbol{p}) = (\boldsymbol{x}-\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a})\kappa^2$$

or, putting $\rho^2 = 1/\kappa^2$,

$$(\boldsymbol{x}-\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a}) = \rho^2.$$

This is the equation of a sphere of radius ρ about the point whose coordinate vector is $\boldsymbol{a}$.

8.2.6 – CODAZZI'S EQUATIONS

An important relation exists between the first and the second fundamental tensors. On differentiating

$$h_{\kappa\mu} = \boldsymbol{n}\,\partial_{\kappa\mu}\boldsymbol{x}$$

we get

$$\partial_\lambda h_{\kappa\mu} = \partial_\lambda \boldsymbol{n}\,\partial_{\kappa\mu}\boldsymbol{x} + \boldsymbol{n}\,\partial_{\kappa\mu\lambda}\boldsymbol{x}.$$

A direct consequence of this result is

$$\partial_\lambda h_{\kappa\mu} - \partial_\mu h_{\kappa\lambda} = \partial_\lambda \boldsymbol{n}\,\partial_{\kappa\mu}\boldsymbol{x} - \partial_\mu \boldsymbol{n}\,\partial_{\kappa\lambda}\boldsymbol{x}. \tag{8.2–22}$$

Next we observe that we may replace the second derivatives of $\boldsymbol{x}$ by the vectors obtained on applying the operator P, being the projection operator associated with the tangent hyperplane at the given point. Hence we may write (8.2–22) as

$$\partial_\lambda h_{\kappa\mu} + \partial_\mu \boldsymbol{n} \mathsf{P} \partial_{\kappa\lambda}\boldsymbol{x} = \partial_\mu h_{\kappa\lambda} + \partial_\lambda \boldsymbol{n} \mathsf{P} \partial_{\kappa\mu}\boldsymbol{x}.$$

Taking account of (8.2–8) and (7.3–6) we readily find

$$\boxed{\partial_\lambda h_{\kappa\mu} - \Gamma^\nu{}_{\kappa\lambda} h_{\nu\mu} = \partial_\mu h_{\kappa\lambda} - \Gamma^\nu{}_{\kappa\mu} h_{\nu\lambda}.} \tag{8.2–23}$$

These equations, the *equations of Codazzi*, are the desired relations between the first and the second fundamental tensors. In section 8.4.2 we shall derive another set of relations between these two tensors.

8.3 – Spherical representation

8.3.1 – DEFINITION

At every point $\boldsymbol{x}(q^\kappa)$ of a hypersurface there is a unit normal $\boldsymbol{n}(q^\kappa)$, which may be considered as the coordinate vector of a point on a unit hypersphere around the origin of the coordinate system in space. Thus we obtain a mapping of the hypersurface on the hypersphere.

The images of the points on the hypersurface constitute a hypersurface on the sphere if and only if the vectors $\partial_\kappa \boldsymbol{n}$, $\kappa = 1, \ldots, n$, are linearly independent. In other words: when the third fundamental tensor is not singular. In this case the correspondence will be referred to as a *spherical representation* of the hypersurface. The third fundamental tensor is obviously the metric tensor of the spherical image.

Referring to sections 5.3.4 and 8.2.3 we may state:

The principal distortions of the spherical representation are the squares of the principal curvatures.

A direct corollary is:

The spherical representation is conformal if and only if the squares of the principal curvatures are equal.

8.3.2 – The angles between corresponding curves

Let $q^\kappa(s)$, $\kappa = 1, \ldots, n$, denote a curve on the given hypersurface referred to its arc length as parameter. It determines a curve $\boldsymbol{n}(s)$ on the unit hypersphere. In terms of the components $u^\kappa = \mathrm{d}_s q^\kappa$, the norm of the vector $\mathrm{d}_s \boldsymbol{n}$, being a tangent vector of the spherical image of the given curve is

$$\mathrm{d}_s \boldsymbol{n}\, \mathrm{d}_s \boldsymbol{n} = \kappa^2 + \tau^2 \tag{8.3–1}$$

(compare (8.2–13)). Hence the angle ϑ between the tangents at a point of the given curve and at the corresponding point of its image is given by

$$\cos \vartheta = \frac{\mathrm{d}_s \boldsymbol{x}\, \mathrm{d}_s \boldsymbol{n}}{\sqrt{\kappa^2 + \tau^2}}. \tag{8.3–2}$$

From (8.1–4) follows that

$$\mathrm{d}_s \boldsymbol{x}\, \mathrm{d}_s \boldsymbol{n} = -\kappa.$$

Hence (8.3–2) takes the simple form

$$\cos \vartheta = \frac{-\kappa}{\sqrt{\kappa^2 + \tau^2}}. \tag{8.3–3}$$

Since $0 \leqq \vartheta \leqq \pi$ we also have

$$\sin \vartheta = \frac{\tau}{\sqrt{\kappa^2 + \tau^2}}, \tag{8.3–4}$$

for we agreed to take $\tau \geqq 0$. As a consequence:

$$\tan \vartheta = -\frac{\tau}{\kappa}. \tag{8.3–5}$$

We wish now to list some special cases. First we observe that $\vartheta = 0$ or $\vartheta = \pi$ only when $\tau = 0$. This means:

The tangents through corresponding points at a curve and its spherical image are parallel if and only if the curve is a line of curvature.

When $\vartheta = \frac{1}{2}\pi$ then $\kappa = 0$. This means that the curve is an asymptotic curve. Hence:

The tangents through corresponding points of a curve and its spherical image are orthogonal if and only if the curve is asymptotic.

8.3.3 – The fundamental equations of spherical representation

It is natural to introduce a new kind of three index symbols which depend on the third fundamental tensor in the same way as the ordinary three index symbols depend on the first fundamental tensor. First we observe that the vectors $\partial_\kappa \boldsymbol{n}$ and $\partial_\kappa \boldsymbol{x}$, $\kappa = 1, \ldots, n$ span the same hyperspace. In accordance with (7.3–4) we now introduce symbols of the second kind by putting

$$\mathsf{P}\partial_{\lambda\mu}\boldsymbol{n} = \partial_\kappa \boldsymbol{n}\,\Gamma^\kappa{}_{\lambda\mu}. \qquad (8.3\text{–}6)$$

Introducing the symbols of the first kind by

$$\Gamma_{\kappa\lambda\mu} = \partial_\kappa \boldsymbol{n}\partial_{\lambda\mu}\boldsymbol{n} \qquad (8.3\text{–}7)$$

we evidently have

$$\Gamma_{\kappa\lambda\mu} = \Gamma^\nu{}_{\lambda\mu}k_{\kappa\nu}.$$

Proceeding as in section 7.3.2 we easily find

$$\Gamma_{\kappa\lambda\mu} = \tfrac{1}{2}(\partial_\lambda k_{\mu\kappa} + \partial_\mu k_{\kappa\lambda} - \partial_\kappa k_{\lambda\mu}). \qquad (8.3\text{–}9)$$

Next we differentiate (8.2–2) and write the result thus obtained in the form

$$\partial_\lambda h_{\kappa\mu} = -\partial_{\kappa\lambda}\boldsymbol{x}\partial_\mu \boldsymbol{n} - \partial_\kappa \boldsymbol{x}\partial_{\mu\lambda}\boldsymbol{n}. \qquad (8.3\text{–}10)$$

From (7.3–4) and (8.2–2) we deduce

$$\partial_{\kappa\lambda}\boldsymbol{x}\partial_\mu \boldsymbol{n} = -\mathit{\Gamma}^\nu{}_{\kappa\lambda}h_{\nu\mu}$$

and from (8.3–6) and (8.2–2) we get

$$\partial_\kappa \boldsymbol{x}\partial_{\mu\lambda}\boldsymbol{n} = -\Gamma^\nu{}_{\mu\lambda}h_{\nu\kappa}.$$

Thus (8.3–10) appears as

$$\partial_\lambda h_{\kappa\mu} = \mathit{\Gamma}^\nu{}_{\kappa\lambda}h_{\nu\mu} + \Gamma^\nu{}_{\mu\lambda}h_{\nu\kappa}. \qquad (8.3\text{–}11)$$

In these equations we may interchange μ and λ. On subtracting corresponding members we find Codazzi's equations (8.2–23). It should be noticed, however, that in this proof we explicitly made the assumption that the vectors $\partial_\kappa \boldsymbol{n}$ are linearly independent.

The equations (8.3–11) may also be written in the form

$$\partial_\lambda h_{\mu\kappa} = \Gamma^\nu{}_{\mu\lambda}h_{\nu\kappa} + \Gamma^\nu{}_{\kappa\lambda}h_{\nu\mu} \qquad (8.3\text{–}12)$$

and after interchanging μ and λ we find, on subtracting corresponding members:

$$\partial_\lambda h_{\kappa\mu} - \Gamma^\nu{}_{\kappa\lambda}h_{\nu\mu} = \partial_\mu h_{\kappa\lambda} - \Gamma^\nu{}_{\kappa\mu}h_{\nu\lambda}. \qquad (8.3\text{–}13)$$

These are Codazzi's equations for spherical representation.

8.4 – The "Theorema egregium"

8.4.1 – The equations of Gauss

Every vector of the system $\partial_{\lambda\mu}\boldsymbol{x}$ can be decomposed into its projection $\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x}$ on the tangent hyperplane and a vector $\boldsymbol{x}_{\lambda\mu}$ normal to it, i.e.,

$$\partial_{\lambda\mu}\boldsymbol{x} = \mathsf{P}\partial_{\lambda\mu}\boldsymbol{x}+\boldsymbol{x}_{\lambda\mu}. \tag{8.4–1}$$

On multiplying both members by $\boldsymbol{n}$ we find from (8.2–3) that

$$\boldsymbol{x}_{\lambda\mu} = \boldsymbol{n}h_{\lambda\mu}. \tag{8.4–2}$$

Combining this result with (7.3–4) we may write (8.4–1) in the form

$$\boxed{\partial_{\lambda\mu}\boldsymbol{x} = \partial_{\kappa}\boldsymbol{x}\Gamma^{\kappa}{}_{\lambda\mu}+\boldsymbol{n}h_{\lambda\mu}.} \tag{8.4–3}$$

These equations are known as the *equations of Gauss.*

8.4.2 – The fundamental formula

In this section we wish to establish a theorem of fundamental importance. Our starting point will be the formula (8.4–3) which we shall write in the form

$$\partial_{\beta\mu}\boldsymbol{x} = \partial_{\nu}\boldsymbol{x}\Gamma^{\nu}{}_{\beta\mu}+\boldsymbol{n}h_{\beta\mu}. \tag{8.4–4}$$

By differentiating we get

$$\partial_{\beta\mu\lambda}\boldsymbol{x} = \partial_{\nu\lambda}\boldsymbol{x}\Gamma^{\nu}{}_{\beta\mu}+\boldsymbol{x}_{\kappa}\partial_{\lambda}\Gamma^{\kappa}{}_{\beta\mu}+\partial_{\lambda}\boldsymbol{n}\,h_{\beta\mu}+\boldsymbol{n}\partial_{\lambda}h_{\beta\mu}. \tag{8.4–5}$$

Next we multiply both members by $\boldsymbol{x}_{\alpha}$ and, taking account of (7.3–5), (7.3–7) and (8.2–2), we get

$$\boldsymbol{x}_{\alpha}\,\partial_{\beta\mu\lambda}\boldsymbol{x} = g_{\alpha\kappa}\Gamma^{\kappa}{}_{\nu\lambda}\Gamma^{\nu}{}_{\beta\mu}+g_{\alpha\kappa}\partial_{\lambda}\Gamma^{\kappa}{}_{\beta\mu}-h_{\alpha\lambda}h_{\beta\mu}. \tag{8.4–6}$$

Since the expression on the left is symmetric with respect to λ and μ we readily find

$$h_{\alpha\lambda}h_{\beta\mu}-h_{\alpha\mu}h_{\beta\lambda} = g_{\alpha\kappa}R^{\kappa}{}_{\beta\lambda\mu} \tag{8.4–7}$$

with

$$R^{\kappa}{}_{\beta\lambda\mu} = \partial_{\lambda}\Gamma^{\kappa}{}_{\beta\mu}-\partial_{\mu}\Gamma^{\kappa}{}_{\beta\lambda}+\Gamma^{\kappa}{}_{\nu\lambda}\Gamma^{\nu}{}_{\beta\mu}-\Gamma^{\kappa}{}_{\nu\mu}\Gamma^{\nu}{}_{\beta\lambda}, \tag{8.4–8}$$

or symbolically:

$$R^{\kappa}{}_{\beta\lambda\mu} = \begin{vmatrix} \partial_{\lambda} & \partial_{\mu} \\ \Gamma^{\kappa}{}_{\beta\lambda} & \Gamma^{\kappa}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Gamma^{\kappa}{}_{\nu\lambda} & \Gamma^{\kappa}{}_{\nu\mu} \\ \Gamma^{\nu}{}_{\beta\lambda} & \Gamma^{\nu}{}_{\beta\mu} \end{vmatrix}. \tag{8.4–9}$$

Writing $R_{\alpha\beta\lambda\mu}$ instead of the right-hand member of (8.4–7) we also have

$$\boxed{h_{\alpha\lambda}h_{\beta\mu}-h_{\alpha\mu}h_{\beta\lambda} = R_{\alpha\beta\lambda\mu}.} \tag{8.4–10}$$

This result is very remarkable. In fact, the tensor with components $R_{\alpha\beta\lambda\mu}$ depends on the metric tensor only, and (8.4–10) expresses the fact that it is possible to write down an expression in terms of the components of the second fundamental tensor, which is an isometric invariant. The tensor $R_{\alpha\beta\lambda\mu}$ is named *Riemann's tensor* and will be studied thoroughly in the next chapter.

We wish to conclude this section by deriving some other consequences from (8.4–5). On multiplying both members by $\boldsymbol{n}$ we get

$$\boldsymbol{n}\partial_{\beta\lambda\mu}\boldsymbol{x} = h_{\nu\lambda}\Gamma^{\nu}{}_{\beta\mu}+\partial_{\lambda}h_{\beta\mu}.$$

Hence, since the expression on the left is symmetric with respect to λ and μ

$$\partial_{\lambda}h_{\beta\mu}-\Gamma^{\nu}{}_{\beta\lambda}h_{\nu\mu} = \partial_{\mu}h_{\beta\lambda}-\Gamma^{\nu}{}_{\beta\mu}h_{\nu\lambda},$$

the equations of Codazzi.

On transvecting by $g^{\alpha\beta}$ both members of (8.4–6) we readily find, taking account of (8.2–14),

$$k_{\lambda\mu} = \partial_{\lambda}\Gamma^{\beta}{}_{\beta\mu}+\Gamma^{\beta}{}_{\nu\lambda}\Gamma^{\nu}{}_{\beta\mu}-\boldsymbol{x}^{\beta}\,\partial_{\beta\mu\lambda}\boldsymbol{x}. \tag{8.4–11}$$

8.4.3 – The "Theorema egregium"

A more geometric interpretation of the main result obtained in the previous section may be found when we recall the formulas (3.3–6) and (3.3–8) and apply them to the second fundamental tensor. Thus we arrive at

$$\sum_{h<k}\underset{h}{\kappa}\,\underset{k}{\kappa} = -\tfrac{1}{2}g^{\alpha\mu}g^{\beta\lambda}\begin{vmatrix} h_{\alpha\lambda} & h_{\alpha\mu} \\ h_{\beta\lambda} & h_{\beta\mu}\end{vmatrix} = -\tfrac{1}{2}g^{\alpha\mu}g^{\beta\lambda}R_{\alpha\beta\lambda\mu}. \tag{8.4–12}$$

The scalar

$$R = g^{\alpha\lambda}g^{\beta\mu}R_{\alpha\beta\lambda\mu} \tag{8.4–13}$$

is called the *scalar curvature* of the hypersurface. Let K denote the mean contribution of each term in the left-hand member of (8.4–12) to the sum. Then we may write this sum as

$$\sum_{h<k}\underset{h}{\kappa}\,\underset{k}{\kappa} = \tfrac{1}{2}n(n-1)K \tag{8.4–14}$$

and we also have

$$R = n(1-n)K. \tag{8.4–15}$$

The scalar K is called the *Gaussian curvature*. Our result can now be stated in the following form:

The Gaussian curvature is an invariant for isometric mapping.

This famous theorem was discovered by Gauss for the case of a surface in a three dimensional space. The importance of the theorem is expressed by the name "Theorema egregium".

Finally we wish to evaluate Riemann's tensor for a hypersphere, making use of the results of section 8.1.2. Differentiating (8.1–11) we get

$$\partial_\lambda \boldsymbol{n} = - \partial_\lambda \boldsymbol{x} \, (1/\rho),$$

whence

$$\partial_\lambda \boldsymbol{n} \, \partial_\mu \boldsymbol{x} = - \partial_\lambda \boldsymbol{x} \partial_\mu \boldsymbol{x}/\rho$$

or

$$h_{\lambda\mu} = \frac{1}{\rho} g_{\lambda\mu}.$$

Thus we arrive at

$$R_{\alpha\beta\lambda\mu} = \frac{1}{\rho^2} (g_{\alpha\lambda} g_{\beta\mu} - g_{\alpha\mu} g_{\beta\lambda}). \tag{8.4–16}$$

The scalar curvature is

$$R = \frac{1}{\rho^2} g^{\alpha\mu} g^{\beta\lambda} (g_{\alpha\lambda} g_{\beta\mu} - g_{\alpha\mu} g_{\beta\lambda})$$

$$= \frac{1}{\rho^2} (\delta^\beta_\alpha \delta^\alpha_\beta - n^2) = \frac{1}{\rho^2} n(1-n).$$

Hence the Gaussian curvature is

$$K = \frac{1}{\rho^2}.$$

It is now easy to check (8.4–14), for the principal curvatures are $1/\rho$.

CHAPTER 9

THEORY OF CURVATURE OF GENERAL MANIFOLDS

The theory of curvature of hypersurfaces was much facilitated by the fact that a well-defined field of normal vectors was available. In the case of general manifolds no such simple entity exists, but fortunately, we can derive expressions which depend only on the first fundamental tensor and are therefore marked out for generalization. The most important tensor which lends itself to this purpose is Riemann's tensor, and the discussions in this chapter will centre around this fundamental entity. This tensor also renders services in problems concerning geodesic and conformal correspondence.

The key of the theory lies in the techniques connected with the so-called covariant differentiation. This subject matter will be the starting-point of our discussion.

9.1 – Covariant differentiation

9.1.1 – COVARIANT DIFFERENTIATION OF A VECTOR

Suppose we have a field

$$\boldsymbol{a}(q^{\kappa}), \qquad \kappa = 1, \ldots, n, \tag{9.1–1}$$

of tangential vectors on a manifold of dimension n, ($n \geqq 2$). It is assumed that the manifold lies in a metric space whose dimension is three at least.

By partial differentiation we obtain the set of vectors

$$\partial_{\lambda}\boldsymbol{a}, \qquad \lambda = 1, \ldots, n. \tag{9.1–2}$$

Two such sets corresponding to different frames (κ) and (κ') are related by the equations:

$$\partial_{\lambda'}\boldsymbol{a} = \partial_{\lambda}\boldsymbol{a}\partial_{\lambda'}q^{\lambda}, \qquad \lambda' = 1', \ldots, n'. \tag{9.1–3}$$

Hence the quantities $\partial_{\lambda}\boldsymbol{a}$ are the components of a vectorial tensor of valency one (see section 4.2.4). In general these components are not included in the tangent space of the manifold at the point where they are evaluated. Just as in the case of geodesic differentiation it turns out to be convenient to replace them by their projections on the tangent space, i.e., we may introduce the system

$$\mathsf{D}_\lambda \boldsymbol{a} = \mathsf{P}\partial_\lambda \boldsymbol{a}, \qquad \lambda = 1, \ldots, n, \tag{9.1–4}$$

where P denotes the projection operator associated with the tangent space. The combined process $\mathsf{P}\partial_\lambda$ is called the *covariant differentiation* of the vector $\boldsymbol{a}$.

Since P is linear we also have

$$\mathsf{D}_{\lambda'} \boldsymbol{a} = \mathsf{D}_\lambda \boldsymbol{a} \partial_{\lambda'} q^\lambda, \tag{9.1–5}$$

that is to say, the $\mathsf{D}_\lambda \boldsymbol{a}$ are the components of a vectorial tensor of valency one, the *covariant derivative of* $\boldsymbol{a}$.

It is clear that

$$\mathsf{D}_\lambda (\boldsymbol{a}+\boldsymbol{b}) = \mathsf{D}_\lambda \boldsymbol{a}+\mathsf{D}_\lambda \boldsymbol{b}. \tag{9.1–6}$$

Let φ denote a scalar invariant and $\boldsymbol{a}$ the vector field (9.1–1). On applying the operator P to

$$\partial_\lambda (\boldsymbol{a}\varphi) = \partial_\lambda \boldsymbol{a}\, \varphi+\boldsymbol{a}\partial_\lambda \varphi$$

we get

$$\mathsf{D}_\lambda (\boldsymbol{a}\varphi) = \mathsf{D}_\lambda \boldsymbol{a}\, \varphi+\boldsymbol{a}\partial_\lambda \varphi. \tag{9.1–7}$$

Finally we wish to prove Leibniz's rule for the differentiation of the inner product of two vectors in terms of their covariant derivatives. From

$$\partial_\lambda (\boldsymbol{ab}) = \boldsymbol{b}\partial_\lambda \boldsymbol{a}+\boldsymbol{a}\partial_\lambda \boldsymbol{b}$$

we find, observing that

$$\begin{aligned} \boldsymbol{b}\partial_\lambda \boldsymbol{a} &= \boldsymbol{b}\mathsf{P}\partial_\lambda \boldsymbol{a} = \boldsymbol{b}\mathsf{D}_\lambda \boldsymbol{a}, \\ \boldsymbol{a}\partial_\lambda \boldsymbol{b} &= \boldsymbol{a}(\mathsf{P}\partial_\lambda \boldsymbol{b}) = \boldsymbol{a}\mathsf{D}_\lambda \boldsymbol{b}, \end{aligned}$$

the result

$$\partial_\lambda (\boldsymbol{ab}) = \boldsymbol{b}\mathsf{D}_\lambda \boldsymbol{a}+\boldsymbol{a}\mathsf{D}_\lambda \boldsymbol{b}. \tag{9.1–8}$$

9.1.2 – The components of the covariant derivative of a vector

Since $\boldsymbol{a}$ is assumed to be a tangential vector we may represent it as

$$\boldsymbol{a} = \boldsymbol{x}_\kappa a^\kappa = \boldsymbol{x}^\kappa a_\kappa. \tag{9.1–9}$$

The components $\mathsf{D}_\lambda \boldsymbol{a}$ are also tangential vectors. Accordingly we may write

$$\mathsf{D}_\lambda \boldsymbol{a} = \boldsymbol{x}_\kappa \nabla_\lambda a^\kappa \tag{9.1–10}$$

and it follows that

$$\nabla_\lambda a^\kappa = \boldsymbol{x}^\kappa \mathsf{D}_\lambda \boldsymbol{a}. \tag{9.1–11}$$

By virtue of the equations of transformation (2.2–22) and (9.1–5) we may infer that

$$\nabla_{\lambda'} a^{\kappa'} = \nabla_\lambda a^\kappa \partial_\kappa q^{\kappa'} \partial_{\lambda'} q^\lambda. \tag{9.1–12}$$

Thus we see that *the expressions* $\nabla_\lambda a^\kappa$ *are the components of a tensor of valency two*. This tensor is said to be obtained by *differentiating covariantly the contravariant components of the vector* $\boldsymbol{a}$.

On the other hand we may also write

$$\mathsf{D}_\lambda \boldsymbol{a} = \boldsymbol{x}^\kappa \nabla_\lambda a_\kappa \tag{9.1–13}$$

with

$$\nabla_\lambda a_\kappa = \boldsymbol{x}_\kappa \mathsf{D}_\lambda \boldsymbol{a}. \tag{9.1–14}$$

In exactly the same way as above we find

$$\nabla_{\lambda'} a_{\kappa'} = \nabla_\lambda a_\kappa \partial_{\kappa'} q^\kappa \partial_{\lambda'} q^\lambda. \tag{9.1–15}$$

The expressions $\nabla_\lambda a_\kappa$ are the components of a tensor and are said to be obtained *by differentiating covariantly the covariant components of the vector* $\boldsymbol{a}$.

Taking account of (2.2–6) and (2.2–8) we find from (9.1–10) and (9.1–14) that

$$\nabla_\lambda a_\kappa = g_{\kappa\nu} \nabla_\lambda a^\nu, \qquad \nabla_\lambda a^\kappa = g^{\kappa\nu} \nabla_\lambda a_\nu,$$

or alternatively

$$\nabla_\lambda (g_{\kappa\nu} a^\nu) = g_{\kappa\nu} \nabla_\lambda a^\nu, \qquad \nabla_\lambda (g^{\kappa\nu} a_\nu) = g^{\kappa\nu} \nabla_\lambda a_\nu. \tag{9.1–16}$$

In words:

The process of lowering and raising the index of the components of a vector and the process of covariant differentiation may be interchanged.

This may also be expressed by saying that *the components of the metric tensor behave like constants in the process of covariant differentiation.* This result will be generalized in section 9.1.7.

On differentiating the second member of (9.1–9) we get

$$\partial_\lambda \boldsymbol{a} = \boldsymbol{x}_\kappa \partial_\lambda a^\kappa + \partial_{\kappa\lambda} \boldsymbol{x}\, a^\kappa$$

and on applying the operator P we have in view of (7.3–4)

$$\mathsf{D}_\lambda \boldsymbol{a} = \boldsymbol{x}_\kappa \partial_\lambda a^\kappa + \boldsymbol{x}_\nu \Gamma^\nu{}_{\kappa\lambda} a^\kappa = \boldsymbol{x}_\kappa (\partial_\lambda a^\kappa + \Gamma^\kappa{}_{\nu\lambda} a^\nu).$$

Hence

$$\boxed{\nabla_\lambda a^\kappa = \partial_\lambda a^\kappa + \Gamma^\kappa{}_{\nu\lambda} a^\nu.} \tag{9.1–17}$$

Again, on differentiating the last member of (9.1–9) we get

$$\partial_\lambda \boldsymbol{a} = \boldsymbol{x}^\kappa \partial_\lambda a_\kappa + \partial_\lambda \boldsymbol{x}^\kappa a_\kappa,$$

whence, by applying the operator P, taking account of (7.5–11),

$$\mathsf{D}_\lambda a = x^\kappa \partial_\lambda a_\kappa - x^\nu \Gamma^\kappa{}_{\nu\lambda} a_\kappa = x^\kappa(\partial_\lambda a_\kappa - \Gamma^\nu{}_{\kappa\lambda} a_\nu).$$

Hence

$$\boxed{\nabla_\lambda a_\kappa = \partial_\lambda a_\kappa - \Gamma^\nu{}_{\kappa\lambda} a_\nu.} \qquad (9.1\text{–}18)$$

It is instructive to verify the covariance of the components (9.1–17) and (9.1–18) by direct computation. First we consider (9.1–18). From

$$\partial_{\lambda'} a_{\kappa'} = \partial_\lambda a_\kappa \partial_{\lambda'} q^\lambda \partial_{\kappa'} q^\kappa + a_\kappa \partial_{\lambda'\kappa'} q^\kappa$$

and (7.4–3) modified as

$$\Gamma^{\nu'}{}_{\kappa'\lambda'} a_{\nu'} = \Gamma^{\nu'}{}_{\kappa'\lambda'} a_\nu \partial_{\nu'} q^\nu = \Gamma^\nu{}_{\kappa\lambda} a_\nu \partial_{\kappa'} q^\kappa \partial_{\lambda'} q^\lambda + a_\nu \partial_{\lambda'\kappa'} q^\nu$$

we find, on subtracting corresponding members, the equation (9.1–15). Next we observe that

$$\partial_{\lambda'} a^{\kappa'} = \partial_\lambda a^\kappa \partial_{\lambda'} q^\lambda \partial_{\kappa'} q^\kappa + a^\kappa \partial_{\lambda\kappa} q^{\kappa'} \partial_{\lambda'} q^\lambda$$

and we use (7.4–3) in the form where the primed and unprimed indices are interchanged. Thus

$$\begin{aligned} \Gamma^\kappa{}_{\nu\lambda} a^\nu &= \Gamma^{\kappa'}{}_{\nu'\lambda'} a^\nu \partial_{\kappa'} q^\kappa \partial_\nu q^{\nu'} \partial_{\lambda'} q^\lambda + a^\nu \partial_{\nu\lambda} q^{\kappa'} \partial_{\kappa'} q^\kappa \\ &= \Gamma^{\kappa'}{}_{\nu'\lambda'} a^{\nu'} \partial_{\kappa'} q^\kappa \partial_{\lambda'} q^\lambda + a^\nu \partial_{\nu\lambda} q^{\kappa'} \partial_{\kappa'} q^\kappa. \end{aligned}$$

Again by subtracting corresponding members we find after a slight modification the equation (9.1–12).

Finally we wish to check (9.1–8) by means of (9.1–17) and (9.1–18)

$$\begin{aligned} (\nabla_\lambda a_\kappa) b^\kappa + a_\kappa (\nabla_\lambda b^\kappa) &= (\partial_\lambda a_\kappa - \Gamma^\nu{}_{\kappa\lambda} a_\nu) b^\kappa \\ + a_\kappa (\partial_\lambda b^\kappa + \Gamma^\kappa{}_{\nu\lambda} b^\nu) &= (\partial_\lambda a_\kappa) b^\kappa + a_\kappa \partial_\lambda b^\kappa \\ + \Gamma^\nu{}_{\kappa\lambda} a_\nu b^\kappa + \Gamma^\kappa{}_{\nu\lambda} a_\kappa b^\nu &= (\partial_\lambda a_\kappa) b^\kappa + a_\kappa \partial_\lambda b^\kappa = \partial_\lambda (a_\kappa b^\kappa). \end{aligned}$$

9.1.3 – The divergence of a vector

The concept of covariant differentiation enables us to write the formula (5.4–24) in a more concise form. In fact, we can prove that

$$\boxed{\operatorname{div} \boldsymbol{v} = \nabla_\lambda v^\lambda.} \qquad (9.1\text{–}19)$$

This follows at once from the definition of covariant differentiation and the equation (7.3–14), for

$$\begin{aligned} \nabla_\lambda v^\lambda &= \partial_\lambda v^\lambda + \Gamma^\lambda{}_{\mu\lambda} v^\mu = \partial_\lambda v^\lambda + \frac{\partial_\mu \sqrt{g}}{\sqrt{g}} v^\mu \\ &= \frac{1}{\sqrt{g}} (\sqrt{g}\, \partial_\lambda v^\lambda + \partial_\lambda \sqrt{g}\, v^\lambda) = \frac{1}{\sqrt{g}} \partial_\lambda (\sqrt{g}\, v^\lambda). \end{aligned}$$

Replacing v^λ by $g^{\lambda\mu}\varphi_\mu$ we find, taking account of (9.1–16),

$$\operatorname{div}\operatorname{grad}\varphi = \nabla_\lambda g^{\lambda\mu}\partial_\mu\varphi = g^{\lambda\mu}\nabla_\lambda\partial_\mu\varphi$$

or

$$\boxed{\operatorname{div}\operatorname{grad}\varphi = g^{\lambda\mu}(\partial_{\lambda\mu}\varphi - \Gamma^\kappa{}_{\lambda\mu}\partial_\kappa\varphi).} \tag{9.1–20}$$

The method of obtaining these formulae shows at once that div $\boldsymbol{v}$ and div grad φ are scalar invariants.

9.1.4 – The covariant derivative of a tensor

The process of differentiating covariantly the components of a vector can be extended to the case of an arbitrary tensor. The general case is a little complicated, but we can show the pattern in a simple example.

We consider a tensor field with components

$$a^\kappa{}_{\lambda\mu}. \tag{9.1–21}$$

Next we introduce arbitrary vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ and we form the scalar invariant

$$\varphi = a^\kappa{}_{\lambda\mu}u_\kappa v^\lambda w^\mu. \tag{9.1–22}$$

The partial derivatives of φ with respect to the variables q^ρ, $\rho = 1, \ldots, n$ are the components of a vector, viz.

$$\partial_\rho\varphi = \partial_\rho a^\kappa{}_{\lambda\mu}u_\kappa v^\lambda w^\mu + a^\kappa{}_{\lambda\mu}\partial_\rho u_\kappa v^\lambda w^\mu + a^\kappa{}_{\lambda\mu}u_\kappa\partial_\rho v^\lambda w^\mu + \\ + a^\kappa{}_{\lambda\mu}u_\kappa v^\lambda\partial_\rho w^\mu. \tag{9.1–23}$$

Next we shall write formally

$$\partial_\rho\varphi = \nabla_\rho a^\kappa{}_{\lambda\mu}u_\kappa v^\lambda w^\mu + a^\kappa{}_{\lambda\mu}\nabla_\rho u_\kappa v^\lambda w^\mu + a^\kappa{}_{\lambda\mu}u_\kappa\nabla_\rho v^\lambda w^\mu + \\ + a^\kappa{}_{\lambda\mu}u_\kappa v^\lambda\nabla_\rho w^\mu. \tag{9.1–14}$$

It is evident that the still undefined expression $\nabla_\rho a^\kappa{}_{\lambda\mu}$ must be

$$\boxed{\nabla_\rho a^\kappa{}_{\lambda\mu} = \partial_\rho a^\kappa{}_{\lambda\mu} + \Gamma^\kappa{}_{\nu\rho}a^\nu{}_{\lambda\mu} - \Gamma^\nu{}_{\lambda\rho}a^\kappa{}_{\nu\mu} - \Gamma^\nu{}_{\mu\rho}a^\kappa{}_{\lambda\nu}.} \tag{9.1–25}$$

The last three terms occurring on the right of (9.1–24) are components of a vector, like the expressions on the left. Hence the first term on the right also represents the components of a vector and on applying the quotient rule of section 3.4.4 we easily deduce that the $\nabla_\rho a^\kappa{}_{\lambda\mu}$ are the components of a tensor. Thus we see that *this process of covariant differentiation yields again a tensor whose valency exceeds that of the original tensor by one.*

After these preliminaries it is not difficult to describe the general situation.

From a tensor with components

$$a^{\lambda_1 \dots \lambda_h}{}_{\mu_1 \dots \mu_k} \tag{9.1–26}$$

with valency $h+k$ arises a tensor with components

$$\nabla_\rho a^{\lambda_1 \dots \lambda_h}{}_{\mu_1 \dots \mu_k} \tag{9.1–27}$$

with valency $h+k+1$ such that the components (9.1–27) consist of the term

$$\partial_\rho a^{\lambda_1 \dots \lambda_h}{}_{\mu_1 \dots \mu_k} \tag{9.1–28}$$

together with h terms of the type

$$\Gamma^{\lambda_t}{}_{\nu\rho} a^{\dots \nu \dots}{}_{\mu_1 \dots \mu_k}, \tag{9.1–29}$$

where ν takes the place of λ_t, $t = 1, \dots, h$, among the upper indices, and k terms of the type

$$-\Gamma^{\nu}{}_{\mu_t \rho} a^{\lambda_1 \dots \lambda_h}{}_{\dots \nu \dots}, \tag{9.1–30}$$

where ν takes the place of μ_t, $t = 1, \dots, k$, among the lower indices.

In the case of a scalar φ, being a tensor of valency zero, no terms of the type (9.1–29) or (9.1–30) occur. Hence we may identify $\nabla_\rho \varphi$ with $\partial_\rho \varphi$, i.e.,

$$\boxed{\nabla_\rho \varphi = \partial_\rho \varphi.} \tag{9.1–31}$$

9.1.5 – Codazzi's theorem

In section 8.2.5 we obtained Codazzi's theorem for the tensor $h_{\lambda\mu}$ on a hypersurface in the form

$$\partial_\lambda h_{\kappa\mu} - \Gamma^\nu{}_{\kappa\mu} h_{\nu\lambda} = \partial_\mu h_{\kappa\lambda} - \Gamma^\nu{}_{\kappa\mu} h_{\nu\lambda}.$$

Adding on both sides

$$-\Gamma^\nu{}_{\mu\lambda} h_{\kappa\nu} = -\Gamma^\nu{}_{\lambda\mu} h_{\kappa\nu}$$

we get

$$\partial_\lambda h_{\kappa\mu} - \Gamma^\nu{}_{\kappa\lambda} h_{\nu\mu} - \Gamma^\nu{}_{\mu\lambda} h_{\kappa\nu} = \partial_\mu h_{\kappa\lambda} - \Gamma^\nu{}_{\kappa\mu} h_{\nu\lambda} - \Gamma^\nu{}_{\lambda\mu} h_{\kappa\nu}$$

and this may be written concisely as

$$\boxed{\nabla_\lambda h_{\kappa\mu} = \nabla_\mu h_{\kappa\lambda}.} \tag{9.1–32}$$

The tensor $\nabla_\lambda h_{\kappa\mu}$ is known as *Codazzi's tensor.*

9.1.6 – Leibniz's rule

It is apparent that the process of covariant differentiation applied to the components of the sum of two tensors yields the sum of the covariant derivatives of the components of these tensors. A little more attention is required to establish the statement that for covariant differentiation a rule holds which is the analogue of Leibniz's rule for the derivative of a product. Let us consider the tensor product

$$a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q}. \tag{9.1–33}$$

The component with respect to q^ρ of the covariant derivative contains in the first place a term of the type

$$\partial_\rho a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q}, \tag{9.1–34}$$

together with terms of the type

$$\Gamma^{\lambda_t}{}_{\nu\rho} a^{\dots\nu\dots}{}_{\mu_1\dots\mu_k} b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q}, \qquad t = 1, \dots, h, \tag{9.1–35}$$

and terms of the type

$$-\Gamma^{\nu}{}_{\mu_t\rho} a^{\lambda_1\dots\lambda_h}{}_{\dots\nu\dots} b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q}, \qquad t = 1, \dots, k, \tag{9.1–36}$$

where ν runs through the indices of the first factor. All these terms constitute the expression

$$\nabla_\rho a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q}. \tag{9.1–37}$$

Again we have a term

$$a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} \partial_\rho b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q} \tag{9.1–38}$$

together with terms of the type

$$\Gamma^{\alpha_t}{}_{\nu\rho} a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} b^{\dots\nu\dots}{}_{\beta_1\dots\beta_q}, \qquad t = 1, \dots, p, \tag{9.1–39}$$

and terms of the type

$$-\Gamma^{\nu}{}_{\beta_t\rho} a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} b^{\alpha_1\dots\alpha_p}{}_{\dots\nu\dots}, \qquad t = 1, \dots, q, \tag{9.1–40}$$

constituting the expression

$$a^{\lambda_1\dots\lambda_h}{}_{\mu_1\dots\mu_k} \nabla_\rho b^{\alpha_1\dots\alpha_p}{}_{\beta_1\dots\beta_q}. \tag{9.1–41}$$

The expressions (9.1–37) and (9.1–41) involve all terms obtained by differentiating covariantly the components (9.1–33). Omitting the indices we therefore have

$$\nabla_\rho(ab) = \nabla_\rho a\, b + a \nabla_\rho b, \tag{9.1–42}$$

the desired result.

Finally we wish to show, that the rule holds also in the case of transvection.

To this end it is sufficient to show that the processes of contraction and covariant differentiation applied to the components of a tensor may be interchanged. The tensor

$$\nabla_\rho a^{\lambda_1 \ldots \lambda_k}{}_{\mu_1 \ldots \mu_k} \tag{9.1–43}$$

contains the terms

$$\Gamma^{\lambda_p}{}_{\nu\rho}\, a^{\ldots\nu\ldots}{}_{\mu_1\ldots\mu_k} - \Gamma^{\nu}{}_{\mu_q\rho}\, a^{\lambda_1\ldots\lambda_k}{}_{\ldots\nu\ldots},$$

where ν takes the place of λ_p in the first term and that of μ_q in the second term. Contraction on these indices, that is, putting $\lambda_p = \mu_q = \kappa$ and performing the summation with respect to κ, yields

$$\Gamma^{\kappa}{}_{\nu\rho} a^{\ldots\nu\ldots}{}_{\ldots\kappa\ldots} - \Gamma^{\nu}{}_{\kappa\rho} a^{\ldots\kappa\ldots}{}_{\ldots\nu\ldots} = 0.$$

Hence, differentiating first and contracting afterwards yields the same result as differentiating the components of the contracted tensor.

9.1.7 – Ricci's theorem

A very remarkable result is obtained when we apply the process of covariant differentiation to the components of the metric tensor. In view of Leibniz's rule we deduce at once from (9.1–16)

$$\boxed{\nabla_\rho g_{\lambda\mu} = 0, \qquad \nabla_\rho g^{\lambda\mu} = 0.} \tag{9.1–44}$$

Thus we have obtained Ricci's theorem:

The covariant derivatives of the components of the metric tensor are identically zero.

We did not verify this theorem in the case of mixed components. But also in this case the proof does not present any difficulty, for

$$\nabla_\rho a_\lambda = \nabla_\rho (a_\mu \delta^\mu_\lambda) = \delta^\mu_\lambda \nabla_\rho a_\mu + a_\mu \nabla_\rho \delta^\mu_\lambda,$$

whence

$$\boxed{\nabla_\rho \delta^\mu_\lambda = 0.} \tag{9.1–45}$$

It is instructive to derive these results by direct computation. By virtue of (7.3–4) we have

$$\begin{aligned}\partial_\rho g_{\lambda\mu} &= \partial_\rho (\mathbf{x}_\lambda \mathbf{x}_\mu) = (\partial_{\lambda\rho}\mathbf{x})\mathbf{x}_\mu + \mathbf{x}_\lambda \partial_{\mu\rho}\mathbf{x} = (\mathsf{P}\partial_{\lambda\rho}\mathbf{x})\mathbf{x}_\mu + \mathbf{x}_\lambda\, \mathsf{P}\partial_{\mu\rho}\mathbf{x} \\ &= \Gamma^\nu{}_{\lambda\rho}\mathbf{x}_\nu \mathbf{x}_\mu + \Gamma^\nu{}_{\mu\rho}\mathbf{x}_\lambda \mathbf{x}_\nu = \Gamma^\nu{}_{\lambda\rho} g_{\nu\mu} + \Gamma^\nu{}_{\mu\rho} g_{\lambda\nu}.\end{aligned}$$

This proves the first equation (9.1–44). Again, taking account of (7.5–11):

$$\begin{aligned}\partial_\rho g^{\lambda\mu} &= \partial_\rho (\mathbf{x}^\lambda \mathbf{x}^\mu) = (\partial_\rho \mathbf{x}^\lambda)\mathbf{x}^\mu + \mathbf{x}^\lambda\, \partial_\rho \mathbf{x}^\mu = (\mathsf{P}\partial_\rho \mathbf{x}^\lambda)\mathbf{x}^\mu + \mathbf{x}^\lambda\, \mathsf{P}\partial_\rho \mathbf{x}^\mu \\ &= -\Gamma^\lambda{}_{\nu\rho}\mathbf{x}^\nu \mathbf{x}^\mu - \Gamma^\mu{}_{\nu\rho}\mathbf{x}^\lambda \mathbf{x}^\nu = -\Gamma^\lambda{}_{\nu\rho} g^{\nu\mu} - \Gamma^\mu{}_{\nu\rho} g^{\lambda\nu}.\end{aligned}$$

This proves the second equation (9.1–44). Finally

$$\nabla_\rho \delta^\mu_\lambda = \partial_\rho \delta^\mu_\lambda + \Gamma^\mu{}_{\nu\rho}\delta^\nu_\lambda - \Gamma^\nu{}_{\lambda\rho}\delta^\mu_\nu = \Gamma^\mu{}_{\lambda\rho} - \Gamma^\mu{}_{\lambda\rho} = 0.$$

Thus we have also checked (9.1–45).

An important consequence of Ricci's theorem is that also for the components of a general tensor the process of lowering and raising the indices is reversible with covariant differentiation. Hence the different types of components representing the same tensor yield components which also represent the same tensor. This latter tensor can therefore be called the *covariant derivative* of the original tensor.

9.2 – The tensor of Riemann

9.2.1 – The normal vector tensor

The vector $\boldsymbol{x}_{\lambda\mu}$ obtained from $\partial_{\lambda\mu}\boldsymbol{x}$ by subtracting from it its projection $\mathsf{P}\partial_{\lambda\mu}\boldsymbol{x}$ on the tangent space is orthogonal to the manifold. In view of (7.3–4) we may write

$$\boldsymbol{x}_{\lambda\mu} = \partial_\mu \boldsymbol{x}_\lambda - \boldsymbol{x}_\kappa \Gamma^\kappa{}_{\lambda\mu}. \tag{9.2–1}$$

The expression on the right looks like a covariant derivative, that is to say

$$\boldsymbol{x}_{\lambda\mu} = \nabla_\mu \boldsymbol{x}_\lambda,$$

for the process of formal covariant differentiation may also be applied to the components of a vector tensor. As a consequence we may expect that the $\boldsymbol{x}_{\lambda\mu}$ obey the rule of transformation

$$\boldsymbol{x}_{\lambda'\mu'} = \boldsymbol{x}_{\lambda\mu}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu. \tag{9.2–2}$$

This may also be verified by straightforward computation. In fact, it follows from

$$\partial_{\lambda'} \boldsymbol{x} = \partial_\lambda \boldsymbol{x}\, \partial_{\lambda'} q^\lambda$$

that

$$\partial_{\lambda'\mu'} \boldsymbol{x} = \partial_{\lambda\mu} \boldsymbol{x}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + \partial_\lambda \boldsymbol{x}\, \partial_{\lambda'\mu'} q^\lambda$$

and

$$\mathsf{P}\partial_{\lambda'\mu'} \boldsymbol{x} = \mathsf{P}\partial_{\lambda\mu} \boldsymbol{x}\, \partial_{\lambda'} q^\lambda\, \partial_{\mu'} q^\mu + \partial_\lambda \boldsymbol{x}\, \partial_{\lambda'\mu'} q^\lambda.$$

Subtracting corresponding members of these equations yields at once (9.2–2).

9.2.2 – The tensor of Riemann

In section 8.4.2 we obtained a fundamental relation between the first and second fundamental tensors expressed by (8.4–10). In view of (8.4–12) this equation may be cast into the form

$$R_{\alpha\beta\lambda\mu} = x_{\alpha\lambda}x_{\beta\mu} - x_{\alpha\mu}x_{\beta\lambda}. \tag{9.2-3}$$

The expression on the right has significance for general manifolds. Hence (9.2–3) may serve to define the *tensor of Riemann* for an arbitrary manifold of dimension $\geqq 2$.

It is immediately seen that this tensor is alternating with respect to the first two and the last two subscripts, i.e.,

$$R_{\alpha\beta\lambda\mu} = -R_{\beta\alpha\lambda\mu} = -R_{\alpha\beta\mu\lambda} = R_{\beta\alpha\mu\lambda}. \tag{9.2-4}$$

It is also possible to interchange the first pair and the second pair of indices:

$$R_{\alpha\beta\lambda\mu} = R_{\lambda\mu\alpha\beta}. \tag{9.2-5}$$

This is a consequence of the symmetry of $x_{\lambda\mu}$, for

$$x_{\alpha\lambda}x_{\beta\mu} - x_{\alpha\mu}x_{\beta\lambda} = x_{\lambda\alpha}x_{\mu\beta} - x_{\lambda\beta}x_{\mu\alpha}.$$

Finally we deduce from

$$\begin{aligned} R_{\alpha\lambda\mu\nu} &= x_{\alpha\mu}x_{\lambda\nu} - x_{\alpha\nu}x_{\lambda\mu}, \\ R_{\alpha\mu\nu\lambda} &= x_{\alpha\nu}x_{\mu\lambda} - x_{\alpha\lambda}x_{\mu\nu}, \\ R_{\alpha\nu\lambda\mu} &= x_{\alpha\lambda}x_{\nu\mu} - x_{\alpha\mu}x_{\nu\lambda}, \end{aligned}$$

and the symmetry of $x_{\lambda\mu}$ the identity

$$R_{\alpha\lambda\mu\nu} + R_{\alpha\mu\nu\lambda} + R_{\alpha\nu\lambda\mu} = 0. \tag{9.2-6}$$

The number of not necessarily vanishing components of the tensor of Riemann does not exceed

$$\binom{n}{2}^2 = \tfrac{1}{4}n^2(n-1)^2,$$

as may be seen from (9.2–4) and (9.2–5). The number of relations (9.2–6) is

$$n\binom{n}{3} = \tfrac{1}{6}n^2(n-1)(n-2).$$

Hence the number of not necessarily vanishing components does not exceed

$$\tfrac{1}{4}(n^4-2n^3+n^2) - \tfrac{1}{6}(n^4-3n^3+2n^2) = \tfrac{1}{12}n^2(n^2-1).$$

There are no more independent relations as follows from the fact that the number of independent components is one in the case of a sphere in ordinary space. Hence the number of not necessarily vanishing components is exactly

$$N = \tfrac{1}{12} n^2(n^2-1). \tag{9.2–7}$$

In the case of a hypersphere in an $(n+1)$-dimensional space it is an easy problem to evaluate Riemann's tensor, as we already pointed out in section 8.4.3.

9.2.3 – The identity of Bianchi

We have already pointed out, that the vectorial tensor with components $x_{\lambda\mu}$ is obtained from x_λ by means of covariant formal differentiation. On applying this process again we get a vectorial tensor of valency three,

$$x_{\beta\lambda\mu} = \nabla_\mu x_{\beta\lambda} = \partial_\mu x_{\beta\lambda} - x_{\nu\lambda}\Gamma^\nu{}_{\beta\mu} - x_{\beta\nu}\Gamma^\nu{}_{\lambda\mu}. \tag{9.2–8}$$

Next we wish to prove that the vectors

$$x_{\beta\lambda\mu} - x_{\beta\mu\lambda} \tag{9.2–9}$$

are tangential vectors. This is immediately clear when we insert the expressions for $x_{\lambda\mu}$ into (9.2–8) and collect all tangential vectors into $y_{\beta\lambda\mu}$. The result is

$$\partial_{\beta\lambda\mu}x - \partial_{\nu\mu}x\Gamma^\nu{}_{\beta\mu} - \partial_{\nu\lambda}x\Gamma^\nu{}_{\beta\mu} - \partial_{\beta\nu}x\Gamma^\nu{}_{\lambda\mu} + y_{\beta\lambda\mu}.$$

Observing that the sum preceding $y_{\beta\lambda\mu}$ is symmetric with respect to λ and μ we may infer that (9.2–9) equals $y_{\beta\lambda\mu} - y_{\beta\mu\lambda}$ and this concludes the proof of the assertion.

Let us now apply the process of covariant differentiation to the tensor (9.2–3). We get

$$\nabla_\kappa R_{\alpha\beta\lambda\mu} = x_{\alpha\lambda\kappa}x_{\beta\mu} + x_{\alpha\lambda}x_{\beta\mu\kappa} - x_{\alpha\mu\kappa}x_{\beta\lambda} - x_{\alpha\mu}x_{\beta\lambda\kappa}.$$

By cyclic permutation of the indices κ, λ and μ we find two similar expressions. The sum of the right hand numbers turns out to be

$$(x_{\alpha\lambda\kappa} - x_{\alpha\kappa\lambda})x_{\beta\mu} + x_{\alpha\lambda}(x_{\beta\mu\kappa} - x_{\beta\kappa\mu}) + (x_{\alpha\mu\lambda} - x_{\alpha\lambda\mu})x_{\beta\kappa} + x_{\alpha\mu}(x_{\beta\kappa\lambda} - x_{\beta\lambda\kappa})$$
$$+ (x_{\alpha\kappa\mu} - x_{\alpha\mu\kappa})x_{\beta\lambda} + x_{\alpha\kappa}(x_{\beta\lambda\mu} - x_{\beta\mu\lambda}).$$

In view of the fact that all vectors of the type (9.2–9) are tangential we conclude that the above sum is zero. Thus we have proved *Bianchi's identity*:

$$\boxed{\nabla_\kappa R_{\alpha\beta\lambda\mu} + \nabla_\lambda R_{\alpha\beta\mu\kappa} + \nabla_\mu R_{\alpha\beta\kappa\lambda} = 0.} \tag{9.2–10}$$

9.2.4 – The tensors of Ricci and Einstein

By raising the first index of the components of the tensor of Riemann we obtain the mixed components

$$R^{\kappa}{}_{\beta\lambda\mu} = g^{\kappa\alpha} R_{\alpha\beta\lambda\mu}. \qquad (9.2\text{–}11)$$

Contraction on κ and μ yields *Ricci's tensor*

$$\boxed{R_{\beta\lambda} = R^{\kappa}{}_{\beta\lambda\kappa}.} \qquad (9.2\text{–}12)$$

The tensor of Ricci is symmetric.

In fact, in view of (9.2–5) and (9.2–4) we have

$$R_{\beta\lambda} = g^{\alpha\kappa} R_{\alpha\beta\lambda\kappa} = g^{\alpha\kappa} R_{\lambda\kappa\alpha\beta} = g^{\alpha\kappa} R_{\kappa\lambda\beta\alpha} = R_{\lambda\beta}.$$

Contraction on κ and λ does not yield a new tensor, for

$$R^{\kappa}{}_{\beta\kappa\mu} = -R^{\kappa}{}_{\beta\mu\kappa} = -R_{\beta\mu}.$$

Contraction on κ and β produces a tensor which is identically zero, for

$$R^{\kappa}{}_{\kappa\lambda\mu} = g^{\alpha\kappa} R_{\alpha\kappa\lambda\mu} = -g^{\alpha\kappa} R_{\kappa\alpha\lambda\mu} = -R^{\alpha}{}_{\alpha\lambda\mu}.$$

Ricci's tensor, however, is not necessarily zero. This may be checked in the case of a hypersphere. Making use of (8.4–16) we easily find

$$R_{\beta\lambda} = \frac{1}{\rho^2} g^{\alpha\kappa}(g_{\alpha\lambda} g_{\beta\kappa} - g_{\alpha\kappa} g_{\beta\lambda}) = \frac{1}{\rho^2}(\delta^{\alpha}_{\beta} g_{\alpha\lambda} - n g_{\beta\lambda}) = \frac{1}{\rho^2}(1-n) g_{\beta\lambda}.$$

Next we wish to derive an interesting consequence of Bianchi's identity. First we observe that it can be written in the form

$$\nabla_{\kappa} R^{\nu}{}_{\beta\lambda\mu} + \nabla_{\lambda} R^{\nu}{}_{\beta\mu\kappa} + \nabla_{\mu} R^{\nu}{}_{\beta\kappa\lambda} = 0. \qquad (9.2\text{–}13)$$

Contraction on ν and μ yields

$$\nabla_{\kappa} R_{\beta\lambda} - \nabla_{\lambda} R_{\beta\kappa} + \nabla_{\mu} R^{\mu}{}_{\beta\kappa\lambda} = 0.$$

Transvecting this result by $g^{\beta\lambda}$ we get (R being the scalar (8.4–13))

$$\nabla_{\kappa} R - \nabla_{\lambda} R^{\lambda}_{\kappa} - \nabla_{\mu} R^{\mu}_{\kappa} = 0,$$

for

$$g^{\beta\lambda} R^{\mu}{}_{\beta\kappa\lambda} = g^{\beta\lambda} g^{\alpha\mu} R_{\alpha\beta\kappa\lambda} = -g^{\beta\lambda} g^{\alpha\mu} R_{\beta\alpha\kappa\lambda} = -g^{\alpha\mu} R_{\alpha\kappa} = -R^{\mu}_{\kappa}.$$

Thus we arrive at

$$\nabla_{\nu} R^{\nu}_{\kappa} - \tfrac{1}{2} \nabla_{\kappa} R = 0, \qquad (9.2\text{–}14)$$

or

$$\nabla_\nu (R^\nu_\kappa - \tfrac{1}{2}\delta^\nu_\kappa R) = 0. \tag{9.2–15}$$

By introducing the *tensor of Einstein*

$$\boxed{G^\nu_\kappa = R^\nu_\kappa - \tfrac{1}{2}\delta^\nu_\kappa R,} \tag{9.2–16}$$

we have proved that the *divergence* (See also section 9.2.6) of this tensor vanishes:

$$\boxed{\nabla_\nu G^\nu_\kappa = 0.} \tag{9.2–17}$$

This tensor plays an important part in Einstein's theory of gravitation. Its covariant components are

$$G_{\lambda\mu} = R_{\lambda\mu} - \tfrac{1}{2}g_{\lambda\mu} R. \tag{9.2–18}$$

Contracting the mixed components we obtain the scalar

$$G = G^\nu_\nu = \tfrac{1}{2}R(2-n). \tag{9.2–19}$$

More generally we might ask under what circumstances the divergence of a tensor of the type

$$a^\nu_\kappa = R^\nu_\kappa + \delta^\nu_\kappa \varphi$$

where φ is a scalar invariant, is identically zero. In view of (9.2–13) we evidently have

$$0 = \nabla_\nu a^\nu_\kappa = \tfrac{1}{2}\nabla_\kappa R + \nabla_\kappa \varphi = \nabla_\kappa(\tfrac{1}{2}R+\varphi) = \partial_\kappa(\tfrac{1}{2}R+\varphi),$$

where $\kappa = 1, \ldots, n$. Hence $\tfrac{1}{2}R+\varphi$ is constant and

$$a^\nu_\kappa = R^\nu_\kappa - \tfrac{1}{2}\delta^\nu_\kappa (R+c).$$

Thus we see that a^ν_κ is not essentially different from Einstein's tensor.

9.2.5 – Ricci's commutation laws

Riemann's tensor appears in a rather inexpected way when we apply repeatedly the process of covariant differentiation to a vector or a tensor of higher valency. As we shall see presently the order of performing this process is of importance.

Formal covariant differentiation of

$$x_\alpha x_{\beta\mu} = 0 \tag{9.2–19}$$

yields

$$x_{\alpha\lambda} x_{\beta\mu} + x_\alpha x_{\beta\mu\lambda} = 0. \tag{9.2–20}$$

Interchanging the indices λ and μ and subtracting the result obtained from (9.2–20) leads to

$$R_{\alpha\beta\lambda\mu} = \boldsymbol{x}_\alpha(\boldsymbol{x}_{\beta\lambda\mu}-\boldsymbol{x}_{\beta\mu\lambda}) \tag{9.2–21}$$

whence

$$R^\kappa{}_{\beta\lambda\mu} = \boldsymbol{x}^\kappa(\boldsymbol{x}_{\beta\lambda\mu}-\boldsymbol{x}_{\beta\mu\lambda}). \tag{9.2–22}$$

In section 9.2.3 we verified that the vectors $\boldsymbol{x}_{\beta\lambda\mu}-\boldsymbol{x}_{\beta\mu\lambda}$ are tangential. Hence we also have

$$\boxed{\boldsymbol{x}_{\beta\lambda\mu}-\boldsymbol{x}_{\beta\mu\lambda} = \boldsymbol{x}_\nu R^\nu{}_{\beta\lambda\mu}.} \tag{9.2–23}$$

In the case of the hypersurface *this result implies Codazzi's theorem.* In fact, first we have

$$\boldsymbol{x}_{\beta\lambda\mu} = \nabla_\mu \boldsymbol{x}_{\beta\lambda} = \nabla_\mu(\boldsymbol{n}h_{\beta\lambda}) = \partial_\mu \boldsymbol{n}\, h_{\beta\lambda}+\boldsymbol{n}\nabla_\mu h_{\beta\lambda}.$$

Hence (9.2–23) is equivalent to

$$\partial_\mu \boldsymbol{n}\, h_{\beta\lambda}-\partial_\lambda \boldsymbol{n}\, h_{\beta\mu}-\boldsymbol{n}(\nabla_\mu h_{\beta\lambda}-\nabla_\lambda h_{\beta\mu}) = \boldsymbol{x}_\nu R^\nu{}_{\beta\lambda\mu}.$$

Multiplying both members by $\boldsymbol{n}$ we get

$$\nabla_\lambda h_{\beta\mu} = \nabla_\mu h_{\beta\lambda}.$$

As a by-product we find

$$\partial_\mu \boldsymbol{n}\, h_{\beta\lambda}-\partial_\lambda \boldsymbol{n}\, h_{\beta\mu} = \boldsymbol{x}_\nu R^\nu{}_{\beta\lambda\mu}$$

and Weingarten's equations tell us that this result is equivalent to (8.4–10).

Next we consider the vector

$$\boldsymbol{a} = \boldsymbol{x}_\beta a^\beta. \tag{9.2–24}$$

On performing formally twice the process of covariant differentiation, we find

$$\nabla_\mu\nabla_\lambda \boldsymbol{a} = \boldsymbol{x}_{\beta\lambda\mu}a^\beta+\boldsymbol{x}_{\beta\lambda}\nabla_\mu a^\beta+\boldsymbol{x}_{\beta\mu}\nabla_\lambda a^\beta+\boldsymbol{x}_\beta\nabla_\mu\nabla_\lambda a^\beta. \tag{9.2–25}$$

Observing that

$$\nabla_\mu\nabla_\lambda \boldsymbol{a} = \partial_{\lambda\mu}\boldsymbol{a}-\partial_\nu \boldsymbol{a}\Gamma^\nu{}_{\lambda\mu} = \nabla_\lambda\nabla_\mu \boldsymbol{a} \tag{9.2–26}$$

we readily find, taking account of (9.2–23),

$$\begin{aligned}\boldsymbol{o} &= (\boldsymbol{x}_{\beta\lambda\mu}-\boldsymbol{x}_{\beta\mu\lambda})a^\beta+\boldsymbol{x}_\kappa(\nabla_\mu\nabla_\lambda a^\kappa-\nabla_\lambda\nabla_\mu a^\kappa)\\ &= \boldsymbol{x}_\kappa R^\kappa{}_{\beta\lambda\mu}a^\beta+\boldsymbol{x}_\kappa(\nabla_\mu\nabla_\lambda a^\kappa-\nabla_\lambda\nabla_\mu a^\kappa).\end{aligned}$$

Hence, writing ν rather than β,

$$\boxed{\nabla_\lambda\nabla_\mu a^\kappa-\nabla_\mu\nabla_\lambda a^\kappa = R^\kappa{}_{\nu\lambda\mu}a^\nu.} \tag{9.2–27}$$

Lowering the index κ we also have

$$\nabla_\lambda \nabla_\mu a_\kappa - \nabla_\mu \nabla_\lambda a_\kappa = R_{\kappa\nu\lambda\mu} a^\nu = -R_{\nu\kappa\lambda\mu} a^\nu = -R^\nu{}_{\kappa\lambda\mu} a_\nu$$

or, when we write β instead of κ,

$$\boxed{\nabla_\lambda \nabla_\mu a_\beta - \nabla_\mu \nabla_\lambda a_\beta = -R^\nu{}_{\beta\lambda\mu} a_\nu.} \tag{9.2–28}$$

These formulas are known as *Ricci's commutation laws* for covariant differentiation.

It is not difficult to extend these results to the case of a tensor of arbitrary valency. It suffices to show the pattern in a particular example. Suppose, we are given a tensor with components $a^\kappa{}_\beta$. Let $\boldsymbol{v}$, $\boldsymbol{w}$ denote two arbitrary vectors. We form the scalar invariant

$$\varphi = a^\kappa{}_\beta v_\kappa w^\beta. \tag{9.2–29}$$

As in the case of a vectorial invariant (See 9.2–16) we have

$$\nabla_\lambda \nabla_\mu \varphi - \nabla_\mu \nabla_\lambda \varphi = 0.$$

Hence

$$\begin{aligned} 0 &= (\nabla_\lambda \nabla_\mu a^\kappa{}_\beta - \nabla_\mu \nabla_\lambda a^\kappa{}_\beta) v_\kappa w^\nu + a^\nu{}_\beta (\nabla_\lambda \nabla_\mu v_\nu - \nabla_\mu \nabla_\lambda v_\nu) w^\beta + \\ &+ a^\kappa{}_\nu v_\kappa (\nabla_\lambda \nabla_\mu w^\nu - \nabla_\mu \nabla_\lambda w^\nu) = (\nabla_\lambda \nabla_\mu a^\kappa{}_\beta - \nabla_\mu \nabla_\lambda a^\kappa{}_\beta) v_\kappa w^\beta + \\ &- a^\nu{}_\beta R^\kappa{}_{\nu\lambda\mu} v_\kappa w^\beta + a^\kappa{}_\nu R^\nu{}_{\beta\lambda\mu} v_\kappa w^\beta. \end{aligned}$$

As a consequence we have

$$\nabla_\lambda \nabla_\mu a^\kappa{}_\beta - \nabla_\mu \nabla_\lambda a^\kappa{}_\beta = R^\kappa{}_{\nu\lambda\mu} a^\nu{}_\beta - R^\nu{}_{\beta\lambda\mu} a^\kappa{}_\nu. \tag{9.2–30}$$

The general result may now be stated as follows: The tensor

$$\nabla_\lambda \nabla_\mu a^{\kappa_1 \ldots \kappa_p}{}_{\beta_1 \ldots \beta_q} - \nabla_\mu \nabla_\lambda a^{\kappa_1 \ldots \kappa_p}{}_{\beta_1 \ldots \beta_q}$$

is a sum of terms

$$R^{\kappa_h}{}_{\nu\lambda\mu} a^{\ldots\nu\ldots}{}_{\beta_1 \ldots \beta_q}, \qquad h = 1, \ldots, p,$$

where ν runs through the indices $\kappa_1, \ldots, \kappa_p$, and a sum of terms

$$-R^\nu{}_{\beta_k \lambda\mu} a^{\kappa_1 \ldots \kappa_h}{}_{\ldots\nu\ldots}, \qquad k = 1, \ldots, q,$$

where ν runs through the indices $\beta_1, \ldots, \beta_q$.

9.2.6 – The laplacian of a vector

In section 9.1.3 we considered the expression (9.1–20), usually called the laplacian of the scalar invariant φ and denoted by $\triangle\varphi$. It is sometimes convenient to introduce a new operation, *contravariant differentiation*, defined by

$$\nabla^\kappa a = g^{\kappa\lambda} \nabla_\lambda a \tag{9.2–31}$$

where a is an arbitrary tensor. The laplacian of φ may now be written as

$$\Delta\varphi = \nabla^\kappa \nabla_\kappa \varphi, \tag{9.2–32}$$

for $\nabla_\kappa \varphi = \partial_\kappa \varphi$.

It is our intention to define the laplacian for a vector. First we consider the *rotation* or the *curl* of a vector v, being the bivector with components

$$\nabla_\lambda v_\mu - \nabla_\mu v_\lambda = \partial_\lambda v_\mu - \partial_\mu v_\lambda. \tag{9.2–33}$$

This bivector will be denoted by

$$\text{rot } v. \tag{9.2–34}$$

By the *divergence* of a tensor of valency two is understood the vector

$$\nabla^\kappa a_{\lambda\kappa} = \nabla_\kappa a_\lambda{}^\kappa. \tag{9.2–35}$$

We have already encountered this process in section 9.2.4 where we considered the divergence of Einstein's tensor.

Next we wish to turn our attention to the vector div rot v having the components

$$\nabla^\alpha (\nabla_\lambda v_\alpha - \nabla_\alpha v_\lambda). \tag{9.2–36}$$

First we observe that on account of Ricci's commutation rule

$$\begin{aligned} \nabla^\alpha \nabla_\lambda v_\alpha &= g^{\alpha\mu} \nabla_\mu \nabla_\lambda v_\alpha = g^{\alpha\mu} \nabla_\lambda \nabla_\mu v_\alpha + g^{\alpha\mu} R^\nu{}_{\alpha\lambda\mu} v_\nu \\ &= \nabla_\lambda \nabla^\alpha v_\alpha + g^{\alpha\mu} g^{\beta\nu} R_{\beta\alpha\lambda\mu} v_\nu \\ &= \nabla_\lambda \nabla^\alpha v_\alpha - g^{\alpha\mu} g^{\beta\nu} R_{\alpha\beta\lambda\mu} v_\nu = \nabla_\lambda \nabla^\alpha v_\alpha - R_{\beta\lambda} v^\beta. \end{aligned}$$

Further we see that $\nabla_\alpha v^\alpha = \nabla^\alpha v_\alpha$ is a scalar invariant, the divergence (9.1–19) of the vector v. Hence $\nabla_\lambda \nabla^\alpha v_\alpha$ is the gradient of the divergence of v. Introducing the *laplacian of* v as a vector Δv with components

$$\nabla^\kappa \nabla_\kappa v_\lambda + R_{\beta\lambda} v^\beta \tag{9.2–37}$$

we evidently have proved the relation

$$\text{div rot } v = \text{grad div } v - \Delta v. \tag{9.2–38}$$

This is a generalization of a well-known formula of vector analysis in ordinary three dimensional space. Since in that space rot v may be considered as a vector, the left-hand side coincides in that case with rot rot v.

9.3 – The Riemannian curvature

9.3.1 – Definition

The *scalar curvature* R of a manifold is

$$R = R_{\alpha\beta\lambda\mu} g^{\alpha\mu} g^{\beta\lambda} = -R_{\alpha\beta\lambda\mu} g^{\alpha\lambda} g^{\beta\mu}. \tag{9.3–1}$$

At a given point of the manifold we introduce an orthonormal frame $\underset{h}{\mathbf{e}}$, $h = 1, \ldots, n$. Making use of the formula (2.3–18) we may represent R by

$$R = -R_{\alpha\beta\lambda\mu} \sum_{h=1}^{n} \sum_{k=1}^{n} \underset{h}{e^{\alpha}} \underset{k}{e^{\beta}} \underset{h}{e^{\lambda}} \underset{k}{e^{\mu}}. \tag{9.3–2}$$

Thus we see that $-R$ is a sum of terms of the type

$$\underset{hk}{K} = R_{\alpha\beta\lambda\mu} \underset{h}{e^{\alpha}} \underset{k}{e^{\beta}} \underset{h}{e^{\lambda}} \underset{k}{e^{\mu}}. \tag{9.3–3}$$

It is clear that

$$\underset{hk}{K} = \underset{kh}{K}, \qquad \underset{hh}{K} = 0. \tag{9.3–4}$$

Introducing the simple unit bivector with components

$$\underset{hk}{e^{\lambda\mu}} = \tfrac{1}{2}(\underset{h}{e^{\lambda}} \underset{k}{e^{\mu}} - \underset{h}{e^{\mu}} \underset{k}{e^{\lambda}}),$$

and using the same arguments that were used in section 4.3.3 for deriving the formula (4.3–13) we see that

$$\underset{hk}{K} = R_{\alpha\beta\lambda\mu} \underset{hk}{e^{\alpha\beta}} \underset{hk}{e^{\lambda\mu}}. \tag{9.3–5}$$

In this expression we may replace $\underset{hk}{e^{\alpha\beta}}$, $\underset{hk}{e^{\beta\mu}}$ by the components of an arbitrary simple bivector which is non-zero and normalized by dividing it by the square root of its norm. Thus we are led to consider the expression

$$\boxed{K(\mathfrak{f}) = \frac{R_{\alpha\beta\lambda\mu} f^{\alpha\beta} f^{\lambda\mu}}{(g_{\alpha\lambda} g_{\beta\mu} - g_{\alpha\mu} g_{\beta\lambda}) f^{\alpha\beta} f^{\lambda\mu}},} \tag{9.3–6}$$

where $\mathfrak{f}$ symbolizes the vector plane defined by the bivector. $K(\mathfrak{f})$ depends on $\mathfrak{f}$ only, as follows at once from (4.3–10).

The scalar invariant (9.3–5) is called the *Riemannian curvature* of the manifold at the given point with respect to the vector plane $\mathfrak{f}$.

In view of (9.3–5) and (9.3–4) we may cast (9.3–2) into the form

$$-\tfrac{1}{2}R = \tfrac{1}{2}\sum_{h=1}^{n} \sum_{k=1}^{n} \underset{hk}{K} = \sum_{h<k} \underset{hk}{K}. \tag{9.3–7}$$

In the last member of this equation $\frac{1}{2}n(n-1)$ terms appear and the mean contribution of each term to the sum is therefore

$$-\tfrac{1}{2}R/\tfrac{1}{2}n(n-1).$$

It is called the *Gaussian curvature K* at the point under consideration and is related to the scalar curvature R according to

$$\boxed{R = n(1-n)K.} \tag{9.3–8}$$

This formula agrees with (8.4–15).

The case $n = 2$ deserves special mention. All components of the Riemannian tensor are either zero or equal to $\pm R_{1212}$. Since $g_{\alpha\lambda}g_{\beta\mu}-g_{\alpha\mu}g_{\beta\lambda}$ has the same properties of symmetry and antisymmetry as $R_{\alpha\beta\lambda\mu}$, and observing that now in (9.3–7) only one term occurs, we may infer that

$$-\tfrac{1}{2}R = \underset{12}{K} = \frac{R_{1212}\,\underset{1}{e^1}\,\underset{2}{e^1}\,\underset{1}{e^2}\,\underset{2}{e^2}}{(g_{11}g_{22}-g_{12}g_{12})\underset{1}{e^1}\,\underset{2}{e^1}\,\underset{1}{e^2}\,\underset{2}{e^2}} = \frac{R_{1212}}{g_{11}g_{22}-g_{12}g_{12}}.$$

For $n = 2$ we see from (9.3–8) that $R = -2K$. Hence:

In the case of a two dimensional manifold (a surface) the components of Riemann's tensor are

$$R_{\alpha\beta\lambda\mu} = K(g_{\alpha\lambda}g_{\beta\mu}-g_{\alpha\mu}g_{\beta\lambda}), \tag{9.3–9}$$

where K denotes the Gaussian curvature.

9.3.2 – The Riccian curvature

The sum of the terms (9.3–3) for $k = 1, \ldots, n$ equals

$$\sum_{k=1}^{n} \underset{hk}{K} = R_{\alpha\beta\lambda\mu}\underset{h}{e^\alpha}\,\underset{h}{e^\lambda}\,g^{\beta\mu} = -R_{\alpha\lambda}\underset{h}{e^\alpha}\,\underset{h}{e^\lambda}. \tag{9.3–10}$$

The number

$$\check{\kappa} = -R_{\lambda\mu}e^\lambda e^\mu \tag{9.3–11}$$

evaluated for the components of a unit vector $\boldsymbol{e}$ is called the *Riccian curvature* in the direction $\boldsymbol{e}$. Hence (9.3–10) expresses:

The sum of the $n-1$ Riemannian curvatures with respect to the planes spanned by a given direction and $n-1$ other directions forming with it an orthonormal frame is equal to the Riccian curvature in this direction.

As an addition to this theorem we may state:

The sum of the Riccian curvatures in n mutually orthogonal directions equals the negative of the scalar curvature.

In fact,

$$\sum_{h=1}^{n} -R_{\lambda\mu} \underset{h}{e^\lambda} \underset{h}{e^\mu} = -R_{\lambda\mu} g^{\lambda\mu} = -R.$$

The main theorem of section 3.2.3 tells us that the tensor of Ricci, being a symmetric tensor, possesses at each point of the manifold a system of n linearly independent principal directions, the so-called *Riccian principal directions.* The associated Riccian curvatures are called the *Riccian principal curvatures.*

On a hypersurface there is an intimate connection between the Riccian curvatures and the normal curvature in the same direction. In fact, in this case the Riemannian tensor takes the form (8.4–10) and it follows that

$$R_{\beta\lambda} = k_{\beta\lambda} - h_{\beta\lambda} h^\mu_\mu \tag{9.3–12}$$

where the $k_{\beta\lambda}$ are the components of the third fundamental tensor. By virtue of (3.3–4) we have

$$h^\mu_\mu = \sum_{h=1}^{n} \underset{h}{\kappa} = nH, \tag{9.3–13}$$

where H denotes the *mean curvature* of the hypersurface at the given point. It follows from (9.3–12) and (9.3–13) that

$$\check{\kappa} = -R_{\beta\lambda} e^\beta e^\lambda = -k_{\beta\lambda} e^\beta e^\lambda + nHh_{\beta\lambda} e^\beta e^\lambda$$

or, taking account of (8.2–13) and (8.2–1),

$$\check{\kappa} = nH\kappa - (\kappa^2 + \tau^2). \tag{9.3–14}$$

Now let $\underset{h}{u}$, $h = 1, \ldots, n$ denote the principal directions of the second fundamental tensor corresponding to the principal curvatures $\underset{h}{\kappa}$. Then

$$h_{\beta\lambda} \underset{h}{u^\lambda} = \underset{h}{\kappa}\, g_{\beta\lambda} \underset{h}{u^\lambda}, \qquad k_{\beta\lambda} \underset{h}{u^\lambda} = \underset{h}{\kappa^2}\, g_{\beta\lambda} \underset{h}{u^\lambda} \tag{9.3–15}$$

and, therefore, in view of (9.3–12) and (9.3–14),

$$-R_{\beta\lambda} \underset{h}{u^\lambda} = g_{\beta\lambda}(nH\underset{h}{\kappa} - \underset{h}{\kappa^2}) \underset{h}{u^\lambda} = \underset{h}{\check{\kappa}}\, g_{\beta\lambda} \underset{h}{u^\lambda}, \tag{9.3–16}$$

for in the principal directions the torsion vanishes. Thus we see that a principal direction is also a Riccian principal direction. Suppose now that the Riccian principal directions are all different. Then the Riccian principal curvatures are also associated with the second fundamental tensor. In this case the principal directions are uniquely determined and we may state:

If on a hypersurface the Riccian principal curvatures at a point are all different, then the Riccian principal directions coincide with the principal directions of the second fundamental tensor.

9.3.3 – The theorem of Schur

In general the Riemannian curvature at a point of a manifold depends on the vector plane $\mathfrak{f}$ to which it is referred. Suppose now that it is independent of $\mathfrak{f}$. Then we may write

$$(R_{\alpha\beta\lambda\mu}-K(g_{\alpha\lambda}g_{\beta\mu}-g_{\alpha\mu}g_{\beta\lambda}))f^{\alpha\beta}f^{\lambda\mu}=0, \tag{9.3–17}$$

where K is the common value of all $K(\mathfrak{f})$ at the given point.

Next we observe, that, if $a_{\alpha\beta\lambda\mu}$ is any tensor of valency four and

$$a_{\alpha\beta\lambda\mu}f^{\alpha\beta}f^{\lambda\mu}=0 \tag{9.3–18}$$

for every simple bivector $f^{\mu\lambda}$, then

$$a_{\alpha\beta\lambda\mu}+a_{\lambda\mu\alpha\beta}-a_{\beta\alpha\lambda\mu}-a_{\alpha\beta\mu\lambda}=0. \tag{9.3–19}$$

Let $a_{\alpha\beta\lambda\mu}$ denote the factor of $f^{\alpha\beta}f^{\lambda\mu}$ on the left of (9.3–17). Taking account of the symmetries and antisymmetries of this tensor we may infer that $a_{\alpha\beta\lambda\mu}=0$, i.e.,

$$R_{\alpha\beta\lambda\mu}=K(g_{\alpha\lambda}g_{\beta\mu}-g_{\alpha\mu}g_{\beta\lambda}). \tag{9.3–20}$$

If, conversely, (9.3–20) is valid at a point, then $K(\mathfrak{f})$ does not depend on the vector plane $\mathfrak{f}$.

It is easy to interpret the factor K occurring in (9.3–20). Transvecting by $g^{\alpha\mu}$ yields

$$R_{\beta\lambda}=K(1-n)g_{\beta\lambda} \tag{9.3–21}$$

and transvecting again by $g^{\beta\lambda}$ we get

$$R=K(1-n)n. \tag{9.3–22}$$

Hence K is the Gaussian curvature at the given point.

Another consequence is that Einstein's tensor (9.2–15) is represented by

$$G_\kappa^\nu=\delta_\kappa^\nu(1-n)(1-\tfrac{1}{2}n)K. \tag{9.3–23}$$

We proceed by establishing a remarkable theorem due to Schur:

If the Riemannian curvature of a manifold of dimension $n>2$ at each point is the same for every vector plane, it does not vary from point to point.

At each point Einstein's tensor is represented by (9.3–23). Since the divergence of this tensor is identically zero we may infer that

$$0 = \nabla_\nu G^\nu_\kappa = (1-n)(1-\tfrac{1}{2}n)\nabla_\kappa K, \qquad \kappa = 1, \ldots, n,$$

whence, since $n > 2$, $\partial_\nu K = \nabla_\nu K = 0$. This proves the assertion.

A manifold having the property mentioned in the above theorem is called *a manifold of constant Riemannian curvature.*

We wish to answer the following question: Which are the hypersurfaces of constant Riemannian curvature?

In ordinary three dimensional space there are various types of surfaces of constant curvature. The situation is quite simple when the dimension is more than two, for:

The only hypersurfaces of constant Riemannian curvature are the hyperspheres, provided the curvature is not zero, and the dimension not less than three.

Let $\underset{h}{e}$, $h = 1, \ldots, n$, denote the vectors of an orthonormal system of principal directions. The Riemannian curvature with respect to the vector plane spanned by $\underset{h}{e}$ and $\underset{k}{e}$ is given by (9.3–3), which now appears as

$$\underset{hk}{K} = (h_{\alpha\lambda}h_{\beta\mu} - h_{\alpha\mu}h_{\beta\lambda})\underset{h}{e^\alpha}\,\underset{h}{e^\lambda}\,\underset{k}{e^\beta}\,\underset{k}{e^\mu} = \underset{h}{\kappa}\,\underset{k}{\kappa},$$

for two different principal directions are conjugate (section 3.2.1). Assuming $n \geqq 3$, we have by hypothesis

$$\underset{h}{\kappa}\,\underset{k}{\kappa} = \underset{k}{\kappa}\,\underset{m}{\kappa} = \underset{h}{\kappa}\,\underset{m}{\kappa} = K.$$

It follows that $\underset{1}{\kappa} = \ldots = \underset{n}{\kappa}$. Hence $K > 0$. Thus we see that every point of the manifold is an umbilical point and the truth of the assertion stated above follows from the theorem of section 8.2.4.

9.3.4 – The Einstein manifold

If all Riccian principal curvatures at a point of a manifold are equal, then we have a relation of the type

$$R_{\lambda\mu} = \rho g_{\lambda\mu}. \tag{9.3–24}$$

Transvecting by $g^{\lambda\mu}$ yields $R = \rho n$. Hence

$$R_{\lambda\mu} = \frac{R}{n} g_{\lambda\mu}. \tag{9.3–25}$$

A manifold is said to be an *Einstein manifold* when (9.3–25) holds throughout the manifold. About these manifolds we have the following theorem:

The scalar curvature of an Einstein manifold whose dimension exceeds two is constant throughout the manifold.

The Einstein tensor of this manifold is

$$G^{\nu}_{\kappa} = R\left(\frac{1}{n} - \frac{1}{2}\right)\delta^{\nu}_{\kappa}.$$

From (9.2–16) follows

$$0 = \nabla_{\nu} G^{\nu}_{\kappa} = \nabla_{\kappa} R\left(\frac{1}{n} - \frac{1}{2}\right).$$

Hence $\partial_{\kappa} R = 0$, $\kappa = 1, \ldots, n$, and this proves the theorem.

In the case $n = 3$ a stronger result can be obtained. At a given point we take an orthonormal frame $\underset{h}{e}$, $h = 1, \ldots, n$. The Riemannian curvature with respect to the plane spanned by $\underset{h}{e}$ and $\underset{k}{e}$ is given by (9.3–3). If the manifold is an Einstein manifold then (9.3–10) takes the form

$$\sum_{k=1}^{n} \underset{hk}{K} = -\frac{R}{n} g_{\alpha\lambda} \underset{h}{e^{\alpha}} \underset{h}{e^{\lambda}} = -\frac{R}{n} = (n-1)K,$$

where K is the Gaussian curvature. Suppose now that $n = 3$. Then

$$\begin{aligned} \underset{12}{K} + \underset{13}{K} \qquad &= 2K, \\ \underset{12}{K} \qquad + \underset{23}{K} &= 2K, \\ \underset{13}{K} + \underset{23}{K} &= 2K. \end{aligned}$$

It follows that $\underset{12}{K} = K$. Now we observe that $\underset{1}{e}$ and $\underset{2}{e}$ may be chosen in an arbitrary vector plane. Hence $\underset{12}{K}$, being equal to the constant Gaussian curvature K, does not depend on the plane to which it is referred, and as an addition to the above theorem we may state:

A three dimensional Einstein manifold is of constant Riemannian curvature.

Notice that a surface is always an Einstein manifold, but its scalar curvature need not be constant.

9.3.5 – The tensor of Weyl

A necessary and sufficient condition for a manifold of dimension $n > 2$ to be of constant Riemannian curvature is that the tensor of Riemann take the form (9.3–20). Eliminating K from (9.3–20) and (9.3–21) we get

$$R_{\alpha\beta\lambda\mu} = \frac{1}{1-n}(g_{\alpha\lambda}R_{\beta\mu} - g_{\alpha\mu}R_{\beta\lambda}). \tag{9.3–26}$$

Hence on a manifold whose tensor of Riemann is represented by (9.3–20) the tensor

$$\boxed{W_{\alpha\beta\lambda\mu} = R_{\alpha\beta\lambda\mu} + \frac{1}{n-1}(g_{\alpha\lambda}R_{\beta\mu} - g_{\alpha\mu}R_{\beta\lambda})} \tag{9.3–27}$$

vanishes. This tensor will be referred to as the *tensor of Weyl*. It is always zero for a surface. Its importance is based on the theorem:

A necessary and sufficient condition for a manifold of dimension exceeding two to be of constant Riemannian curvature is that the tensor of Weyl vanish identically throughout the manifold.

The sufficiency of the condition stated in the theorem remains to be proved. If (9.3–27) is a zero tensor, then (9.3–26) is valid. Hence, transvecting by $g^{\beta\mu}$,

$$g^{\beta\mu}R_{\alpha\beta\lambda\mu} = -g^{\beta\mu}R_{\beta\alpha\lambda\mu} = -R_{\alpha\lambda} = \frac{1}{1-n}(g_{\alpha\lambda}R - R_{\alpha\lambda}),$$

whence

$$R_{\alpha\lambda} = \frac{R}{n}g_{\beta\lambda}. \tag{9.3–26}$$

Inserting this result into (9.3–26) we get

$$R_{\alpha\beta\lambda\mu} = \frac{R}{n(1-n)}(g_{\alpha\lambda}g_{\beta\mu} - g_{\alpha\mu}g_{\beta\lambda}) = K(g_{\alpha\lambda}g_{\beta\mu} - g_{\alpha\mu}g_{\beta\lambda}).$$

The truth of the statement follows from Schur's theorem.

9.4 – Evaluation of the components of the tensor of Riemann

9.4.1 – The mixed components

When we wish to express the mixed components of the tensor of Riemann in terms of the components of the metric tensor we may take the vectorial tensor $\mathbf{x}_{\beta\lambda\mu}$ as a starting-point. According to the definition we have

$$\mathbf{x}_{\beta\lambda\mu} = \partial_\mu \mathbf{x}_{\beta\lambda} - \mathbf{x}_{\nu\lambda}\Gamma^\nu{}_{\beta\mu} - \mathbf{x}_{\beta\nu}\Gamma^\nu{}_{\lambda\mu} \tag{9.4–1}$$

whence

$$\mathbf{x}^\kappa \mathbf{x}_{\beta\lambda\mu} = \mathbf{x}^\kappa \partial_\mu \mathbf{x}_{\beta\lambda} = \mathbf{x}^\kappa \partial_\mu(\partial_{\beta\lambda}\mathbf{x} - \mathbf{x}_\nu \Gamma^\nu{}_{\beta\lambda})$$
$$= \mathbf{x}^\kappa \partial_{\beta\lambda\mu}\mathbf{x} - \mathbf{x}^\kappa \partial_{\mu\nu}\mathbf{x}\Gamma^\nu{}_{\beta\lambda} - \mathbf{x}^\kappa \mathbf{x}_\nu \partial_\mu \Gamma^\nu{}_{\beta\lambda}.$$

Taking account of (7.3–6) we may infer that

$$\mathbf{x}^\kappa \mathbf{x}_{\beta\lambda\mu} = \mathbf{x}^\kappa \partial_{\beta\lambda\mu}\mathbf{x} - \Gamma^\kappa{}_{\nu\mu}\Gamma^\nu{}_{\beta\lambda} - \partial_\mu \Gamma^\kappa{}_{\beta\lambda}.$$

We now immediately infer from (9.2–22) that

$$\boxed{R^\kappa{}_{\beta\lambda\mu} = \partial_\lambda \Gamma^\kappa{}_{\beta\mu} - \partial_\mu \Gamma^\kappa{}_{\beta\lambda} + \Gamma^\kappa{}_{\nu\lambda}\Gamma^\nu{}_{\beta\mu} - \Gamma^\kappa{}_{\nu\mu}\Gamma^\nu{}_{\beta\lambda}.} \tag{9.4–2}$$

This formula can be put into the symbolical form

$$R^\kappa{}_{\beta\lambda\mu} = \begin{vmatrix} \partial_\lambda & \partial_\mu \\ \Gamma^\kappa{}_{\beta\lambda} & \Gamma^\kappa{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Gamma^\kappa{}_{\nu\lambda} & \Gamma^\kappa{}_{\nu\mu} \\ \Gamma^\nu{}_{\beta\lambda} & \Gamma^\nu{}_{\beta\mu} \end{vmatrix}. \tag{9.4–3}$$

It is important to notice that only the Christoffel symbols of the first kind have been involved. This is the same expression as (8.4–9).

9.4.2 – The covariant components

Proceeding along the same lines as in the previous section we also may derive an expression for the covariant components of Riemann's tensor. Multiplying both members of (9.4–1) by $\mathbf{x}_\alpha$ we obtain

$$\mathbf{x}_\alpha \mathbf{x}_{\beta\lambda\mu} = \mathbf{x}_\alpha \partial_\mu(\partial_{\beta\lambda}\mathbf{x} - \mathbf{x}_\nu \Gamma^\nu{}_{\beta\lambda}) = \mathbf{x}_\alpha \partial_\mu(\partial_{\beta\lambda}\mathbf{x} - \mathbf{x}^\nu \Gamma_{\nu\beta\lambda})$$
$$= \mathbf{x}_\alpha \partial_{\beta\lambda\mu}\mathbf{x} - \mathbf{x}_\alpha \partial_\mu \mathbf{x}^\nu \Gamma_{\nu\beta\lambda} - \mathbf{x}_\alpha \mathbf{x}^\nu \partial_\mu \Gamma_{\nu\beta\lambda}.$$

Hence, by virtue of (7.5–11),

$$\mathbf{x}_\alpha \mathbf{x}_{\beta\lambda\mu} = \mathbf{x}_\alpha \partial_{\beta\lambda\mu}\mathbf{x} + \Gamma^\nu{}_{\alpha\mu}\Gamma_{\nu\beta\lambda} - \partial_\mu \Gamma_{\alpha\beta\lambda}.$$

Thus we may infer from (9.2–21) that

$$R_{\alpha\beta\lambda\mu} = \partial_\lambda \Gamma_{\alpha\beta\mu} - \partial_\mu \Gamma_{\alpha\beta\lambda} + \Gamma^\nu{}_{\alpha\mu}\Gamma_{\nu\beta\lambda} - \Gamma^\nu{}_{\alpha\lambda}\Gamma_{\nu\beta\mu}, \tag{9.4–4}$$

or symbolically

$$R_{\alpha\beta\lambda\mu} = \begin{vmatrix} \partial_\lambda & \partial_\mu \\ \Gamma_{\alpha\beta\lambda} & \Gamma_{\alpha\beta\mu} \end{vmatrix} + \begin{vmatrix} \Gamma^\nu{}_{\alpha\mu} & \Gamma^\nu{}_{\alpha\lambda} \\ \Gamma_{\nu\beta\mu} & \Gamma_{\nu\beta\lambda} \end{vmatrix}. \tag{9.4–5}$$

It is instructive to derive (9.4–4) directly from (9.4–2). First we have

$$R_{\alpha\beta\lambda\mu} = g_{\alpha\kappa}\partial_\lambda \Gamma^\kappa{}_{\beta\mu} - g_{\alpha\kappa}\partial_\mu \Gamma^\kappa{}_{\beta\lambda} + \Gamma_{\alpha\nu\lambda}\Gamma^\nu{}_{\beta\mu} - \Gamma_{\alpha\nu\mu}\Gamma^\nu{}_{\beta\lambda}. \tag{9.4–6}$$

Next we observe that

$$g_{\alpha\kappa}\partial_\lambda \Gamma^\kappa{}_{\beta\mu} = \partial_\lambda \Gamma_{\alpha\beta\mu} - \Gamma^\kappa{}_{\beta\mu}\partial_\lambda g_{\alpha\kappa}.$$

By virtue of (7.3–9) we have

$$\Gamma^{\kappa}{}_{\beta\mu}\partial_{\lambda}g_{\alpha\kappa} = \Gamma^{\kappa}{}_{\beta\mu}\Gamma_{\alpha\kappa\lambda}+\Gamma^{\kappa}{}_{\beta\mu}\Gamma_{\kappa\alpha\lambda} = \Gamma_{\alpha\nu\lambda}\Gamma^{\nu}{}_{\beta\mu}+\Gamma_{\nu\beta\mu}\Gamma^{\nu}{}_{\alpha\lambda}.$$

By simple substitution in (9.4–6) we obtain (9.4–4).

9.4.3 – Transformation of the components of the tensor of Riemann by changing the Christoffel symbols

Let us consider a change of the Christoffel symbols of the following type:

$$\overset{*}{\Gamma}{}^{\kappa}{}_{\lambda\mu} = \Gamma^{\kappa}{}_{\lambda\mu}+\Phi^{\kappa}{}_{\lambda\mu} \tag{9.4–7}$$

where the $\Phi^{\kappa}{}_{\lambda\mu}$ are functions of the parameters $q^1, \ldots, q^n$. Such a change occurs in the case of a correspondence between to manifolds $\boldsymbol{x}(q^{\kappa})$ and $\overset{*}{\boldsymbol{x}}(q^{\kappa})$. In that case the $\Phi^{\kappa}{}_{\lambda\mu}$ are the components of a tensor, as we pointed out in section 7.6.1. However, this property is not needed in the subsequent considerations.

By $\overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu}$ we wish to understand the expression on the right of (9.4–2) where the Christoffel symbols are replaced by the corresponding symbols (9.4–7) with an asterisk. In the case of a correspondence between manifolds the $\overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu}$ are the components of the Riemannian tensor of the manifold $\overset{*}{\boldsymbol{x}}(q^{\kappa})$.

By straightforward computation we easily find

$$\overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu} = \begin{vmatrix} \partial_{\lambda} & \partial_{\mu} \\ \Gamma^{\kappa}{}_{\beta\lambda} & \Gamma^{\kappa}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \partial_{\lambda} & \partial_{\mu} \\ \Phi^{\kappa}{}_{\beta\lambda} & \Phi^{\kappa}{}_{\beta\mu} \end{vmatrix} +$$
$$+\begin{vmatrix} \Gamma^{\kappa}{}_{\nu\lambda} & \Gamma^{\kappa}{}_{\nu\mu} \\ \Gamma^{\nu}{}_{\beta\lambda} & \Gamma^{\nu}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Phi^{\kappa}{}_{\nu\lambda} & \Gamma^{\kappa}{}_{\nu\mu} \\ \Phi^{\nu}{}_{\beta\lambda} & \Gamma^{\nu}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Gamma^{\kappa}{}_{\nu\lambda} & \Phi^{\kappa}{}_{\nu\lambda} \\ \Gamma^{\nu}{}_{\beta\mu} & \Phi^{\nu}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Phi^{\kappa}{}_{\nu\lambda} & \Phi^{\kappa}{}_{\nu\mu} \\ \Phi^{\nu}{}_{\beta\lambda} & \Phi^{\nu}{}_{\beta\mu} \end{vmatrix}.$$

Introducing the symbol $\nabla_{\lambda}\Phi^{\kappa}{}_{\beta\mu}$ by

$$\nabla_{\lambda}\Phi^{\kappa}{}_{\beta\mu} = \partial_{\lambda}\Phi^{\kappa}{}_{\beta\mu}+\Gamma^{\kappa}{}_{\nu\lambda}\Phi^{\nu}{}_{\beta\mu}-\Gamma^{\nu}{}_{\beta\lambda}\Phi^{\kappa}{}_{\nu\mu}-\Gamma^{\nu}{}_{\mu\lambda}\Phi^{\kappa}{}_{\beta\nu}$$

we find by an easy computation that

$$\nabla_{\lambda}\Phi^{\kappa}{}_{\beta\mu}-\nabla_{\mu}\Phi^{\kappa}{}_{\beta\lambda} = \begin{vmatrix} \partial_{\lambda} & \partial_{\mu} \\ \Phi^{\kappa}{}_{\beta\lambda} & \Phi^{\kappa}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Phi^{\kappa}{}_{\nu\lambda} & \Gamma^{\kappa}{}_{\nu\mu} \\ \Phi^{\nu}{}_{\beta\lambda} & \Gamma^{\nu}{}_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \Gamma^{\kappa}{}_{\nu\lambda} & \Phi^{\kappa}{}_{\nu\mu} \\ \Gamma^{\nu}{}_{\beta\lambda} & \Phi^{\nu}{}_{\beta\mu} \end{vmatrix}$$

and thus we may infer that

$$\overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu} = R^{\kappa}{}_{\beta\lambda\mu}+\nabla_{\lambda}\Phi^{\kappa}{}_{\beta\mu}-\nabla_{\mu}\Phi^{\kappa}{}_{\beta\lambda}+\begin{vmatrix} \Phi^{\kappa}{}_{\nu\lambda} & \Phi^{\kappa}{}_{\nu\mu} \\ \Phi^{\nu}{}_{\beta\lambda} & \Phi^{\nu}{}_{\beta\mu} \end{vmatrix}. \tag{9.4–8}$$

These are the desired equations of transformation.

9.5 – Geodesic mapping

9.5.1 – The projective curvature tensor

With reference to section 7.6.9 we recall the relation between two corresponding Christoffel symbols on manifolds in geodesic (also called: projective) correspondence. According to (7.6–9) we have

$$\Phi^{\kappa}{}_{\lambda\mu} = \delta^{\kappa}_{\lambda}\varphi_{\mu}+\delta^{\kappa}_{\mu}\varphi_{\lambda}, \tag{9.5–1}$$

where φ is a scalar invariant and $\varphi_{\kappa} = \partial_{\kappa}\varphi$. Writing

$$\varphi_{\lambda\mu} = \nabla_{\mu}\varphi_{\lambda} = \partial_{\lambda\mu}\varphi-\Gamma^{\kappa}{}_{\lambda\mu}\varphi_{\kappa} = \varphi_{\mu\lambda}$$

we evidently have

$$\nabla_{\lambda}\Phi^{\kappa}{}_{\beta\mu}-\nabla_{\mu}\Phi^{\kappa}{}_{\beta\mu} = \delta^{\kappa}_{\mu}\varphi_{\beta\lambda}-\delta^{\kappa}_{\lambda}\varphi_{\beta\mu}.$$

Further

$$\begin{vmatrix} \Phi^{\kappa}{}_{\nu\lambda} & \Phi^{\kappa}{}_{\nu\mu} \\ \Phi^{\nu}{}_{\beta\lambda} & \Phi^{\nu}{}_{\beta\mu} \end{vmatrix} = \begin{vmatrix} \delta^{\kappa}_{\nu}\varphi_{\lambda}+\delta^{\kappa}_{\lambda}\varphi_{\nu} & \delta^{\kappa}_{\nu}\varphi_{\mu}+\delta^{\kappa}_{\mu}\varphi_{\nu} \\ \delta^{\nu}_{\beta}\varphi_{\lambda}+\delta^{\nu}_{\lambda}\varphi_{\beta} & \delta^{\nu}_{\beta}\varphi_{\mu}+\delta^{\nu}_{\mu}\varphi_{\beta} \end{vmatrix}$$

$$= \begin{vmatrix} \delta^{\kappa}_{\lambda}\varphi_{\nu} & \delta^{\kappa}_{\mu}\varphi_{\nu} \\ \delta^{\nu}_{\lambda}\varphi_{\beta} & \delta^{\nu}_{\beta}\varphi_{\mu} \end{vmatrix} = \delta^{\kappa}_{\lambda}\varphi_{\beta}\varphi_{\mu}-\delta^{\kappa}_{\mu}\varphi_{\beta}\varphi_{\lambda}.$$

Using the abbreviation

$$\psi_{\lambda\mu} = \varphi_{\lambda\mu}-\varphi_{\lambda}\varphi_{\mu} \tag{9.5–2}$$

we evidently have

$$\overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu} = R^{\kappa}{}_{\beta\lambda\mu}+\delta^{\kappa}_{\mu}\psi_{\beta\lambda}-\delta^{\kappa}_{\lambda}\psi_{\beta\mu}. \tag{9.5–3}$$

Contraction on κ and μ yields

$$\overset{*}{R}_{\beta\lambda} = R_{\beta\lambda}+(n-1)\psi_{\beta\lambda}. \tag{9.5–4}$$

By inserting $\psi_{\beta\lambda}$ and $\psi_{\beta\mu}$ obtained from (9.5–3) into (9.5–2) we arrive at

$$\overset{*}{W}{}^{\kappa}{}_{\beta\lambda\mu} = W^{\kappa}{}_{\beta\lambda\mu} \tag{9.5–5}$$

with

$$W^{\kappa}{}_{\beta\lambda\mu} = R^{\kappa}{}_{\beta\lambda\mu}+\frac{1}{n-1}(\delta^{\kappa}_{\lambda}R_{\beta\mu}-\delta^{\kappa}_{\mu}R_{\beta\lambda}). \tag{9.5–6}$$

These, however, are the mixed components of Weyl's tensor (9.3–27).

Summing up we may state:

The mixed components of Weyl's tensor are invariant with respect to geodesic mapping.

Weyl's tensor is also referred to as *the projective curvature tensor.*

We conclude this section by the following remark. The relations (7.6–13) between the projective Thomas symbols and the Christoffel symbols are of the type (7.6–9) when φ_κ is replaced by $-(1/(n+1))\Gamma^\nu{}_{\nu\kappa} = -(1/(n+1))$ $\partial_\kappa \log \sqrt{g}$. In this case they fail to be the components of a gradient, but this fact did not play a part in the above computation. It is easily checked, that the computation may be carried out for this new case and we arrive at

$$W^\kappa{}_{\beta\lambda\mu} = P^\kappa{}_{\beta\lambda\mu} + \frac{1}{n-1}(\delta^\kappa_\lambda P_{\beta\mu} - \delta^\kappa_\mu P_{\beta\lambda}), \tag{9.5–7}$$

where $P^\kappa{}_{\beta\lambda\mu}$ is found from $R^\kappa{}_{\beta\lambda\mu}$ by replacing the Christoffel symbols by the invariant Thomas symbols.

It should be noticed that contraction on κ and μ performed on the components of Weyl's tensor yields a tensor which is identically zero.

9.5.2 – A THEOREM OF BELTRAMI

The following theorem is due to Beltrami for the case of a surface. It states: *The only manifolds whose geodesics correspond to the geodesics of a space of constant Riemannian curvature are manifolds of constant Riemannian curvature.*

If the dimension of the manifold exceeds two, then the theorem is a direct consequence of the theorem of section 9.3.6 and equation (9.5–4). This argument does not apply in the case $n = 2$, for then Weyl's tensor is always zero and fails to provide a test for constant Riemannian curvature. The theorem is true, however, in this case, but its proof requires some attention.

We start with (9.5–3) which now takes the form

$$\overset{*}{R}_{\lambda\mu} = R_{\lambda\mu} + \psi_{\lambda\mu}.$$

Hence

$$\psi_{\lambda\mu} = \overset{*}{R}_{\lambda\mu} - R_{\lambda\mu} = Kg_{\lambda\mu} - \overset{*}{K}\overset{*}{g}_{\lambda\mu} \tag{9.5–8}$$

where K and $\overset{*}{K}$ are the Gaussian curvatures of the surfaces under consideration. Equation (9.5–7) is equivalent to

$$\varphi_{\lambda\mu} = \varphi_\lambda\varphi_\mu + Kg_{\lambda\mu} - \overset{*}{K}\overset{*}{g}_{\lambda\mu}. \tag{9.5–9}$$

Next we suppose that $\overset{*}{K}$ is constant throughout the surface. Differentiating both members of (9.5–8) covariantly we get

$$\nabla_\kappa\varphi_{\lambda\mu} = \varphi_{\lambda\kappa}\varphi_\mu + \varphi_\lambda\varphi_{\mu\kappa} + g_{\lambda\mu}\partial_\kappa K - \overset{*}{K}\nabla_\kappa\overset{*}{g}_{\lambda\mu}.$$

Hence

$$\nabla_\kappa \varphi_{\lambda\mu} - \nabla_\mu \varphi_{\lambda\kappa} = \varphi_{\lambda\kappa}\varphi_\mu - \varphi_{\lambda\mu}\varphi_\kappa + g_{\lambda\mu}\partial_\kappa K - g_{\lambda\kappa}\partial_\mu K - \overset{*}{K}(\nabla_\kappa \overset{*}{g}_{\lambda\mu} - \nabla_\mu \overset{*}{g}_{\lambda\kappa}). \tag{9.5–10}$$

By virtue of (9.5–9) we have

$$\varphi_{\lambda\kappa}\varphi_\mu - \varphi_{\lambda\mu}\varphi_\kappa = K(g_{\lambda\kappa}\varphi_\mu - g_{\lambda\mu}\varphi_\kappa) - \overset{*}{K}(\overset{*}{g}_{\lambda\mu}\varphi_\kappa - \overset{*}{g}_{\lambda\mu}\varphi_\kappa). \tag{9.5–11}$$

The Riccian commutation law (9.2–28) tells us that

$$\begin{aligned} \nabla_\kappa \varphi_{\lambda\mu} - \nabla_\mu \varphi_{\lambda\kappa} &= \nabla_\kappa \nabla_\mu \varphi_\lambda - \nabla_\mu \nabla_\kappa \varphi_\lambda = -R^\nu{}_{\lambda\kappa\mu}\varphi_\nu \\ &= R^\nu{}_{\lambda\mu\kappa}\varphi_\nu = K(\delta^\nu_\mu g_{\lambda\kappa} - \delta^\nu_\kappa g_{\lambda\mu})\varphi_\nu = K(g_{\lambda\kappa}\varphi_\mu - g_{\lambda\mu}\varphi_\kappa). \end{aligned} \tag{9.5–12}$$

In order to bring the last term of (9.5–10) into a more manageable form we make use of the fact, that the covariant derivative of the metric tensor on the second manifold is zero, provided this derivative involves the starred Christoffel symbols. That is to say,

$$\partial_\kappa \overset{*}{g}_{\lambda\mu} = \overset{*}{\Gamma}{}^\nu{}_{\lambda\kappa}\overset{*}{g}_{\nu\mu} + \overset{*}{\Gamma}{}^\nu{}_{\mu\kappa}\overset{*}{g}_{\lambda\kappa}.$$

By virtue of (9.5–1) we may rewrite this equation as

$$\partial_\kappa \overset{*}{g}_{\lambda\mu} = \Gamma^\nu{}_{\lambda\kappa}\overset{*}{g}_{\nu\mu} + \Gamma^\nu{}_{\mu\kappa}\overset{*}{g}_{\nu\lambda} + \delta^\nu_\lambda \overset{*}{g}_{\nu\mu}\varphi_\kappa + \delta^\nu_\kappa \overset{*}{g}_{\lambda\nu}\varphi_\mu = \nabla_\kappa \overset{*}{g}_{\lambda\mu} + \overset{*}{g}_{\lambda\mu}\varphi_\kappa + \overset{*}{g}_{\lambda\kappa}\varphi_\mu.$$

Hence

$$-\overset{*}{K}(\nabla_\kappa \overset{*}{g}_{\lambda\mu} - \nabla_\mu \overset{*}{g}_{\lambda\mu}) = \overset{*}{K}(\overset{*}{g}_{\lambda\kappa}\varphi_\mu - \overset{*}{g}_{\lambda\mu}\varphi_\kappa). \tag{9.5–13}$$

Inserting (9.5–11), (9.5–12) and (9.5–13) into (9.5–10) we obtain

$$g_{\lambda\mu}\partial_\kappa K - g_{\lambda\kappa}\partial_\mu K = 0,$$

whence, on transvecting by $g^{\lambda\mu}$

$$2\partial_\kappa K - \partial_\kappa K = \partial_\kappa K = 0, \qquad \kappa = 1, 2.$$

This proves the assertion.

9.6 – Conformal correspondence

9.6.1 – The conformal curvature tensor

A well-known theorem of elementary differential geometry states that it is always possible to bring two surfaces in a three dimensional space into conformal correspondence. The same arguments apply to surfaces in higher dimensional space.

The situation is not quite so simple in the case of manifolds of dimension exceeding two. However, we are able to derive an important conformal invariant which is closely related to the tensor of Riemann.

We recall the formula (7.6–22):

$$\Phi^{\kappa}{}_{\lambda\mu} = \delta^{\kappa}_{\lambda}\chi_{\mu}+\delta^{\kappa}_{\mu}\chi_{\lambda}-g^{\kappa\rho}g_{\lambda\mu}\chi_{\rho}, \tag{9.6–1}$$

where χ is a function of q^{κ}, $\kappa = 1, \ldots, n$, and χ_{κ} stands for $\partial_{\kappa}\chi$. The components of the metric tensors on the two manifolds are related as follows:

$$\overset{*}{g}_{\lambda\mu} = e^{2\chi}g_{\lambda\mu} \tag{9.6–2}$$

$$\overset{*}{g}{}^{\lambda\mu} = e^{-2\chi}g^{\lambda\mu}. \tag{9.6–3}$$

We insert the expressions (9.6–1) into (9.4–8). First we have

$$\nabla_{\lambda}\Phi^{\kappa}{}_{\beta\mu}-\nabla_{\mu}\Phi^{\kappa}{}_{\beta\lambda} = \delta^{\kappa}_{\mu}\chi_{\beta\lambda}-\delta^{\kappa}_{\lambda}\chi_{\beta\mu}+g^{\kappa\rho}(g_{\beta\lambda}\chi_{\rho\mu}-g_{\beta\mu}\chi_{\rho\lambda}), \tag{9.6–4}$$

where

$$\chi_{\lambda\mu} = \nabla_{\mu}\chi_{\lambda} = \partial_{\lambda\mu}\chi-\Gamma^{\kappa}{}_{\lambda\mu}\chi_{\kappa}.$$

Next

$$\begin{vmatrix} \Phi^{\kappa}{}_{\nu\lambda} & \Phi^{\kappa}{}_{\nu\mu} \\ \Phi^{\nu}{}_{\beta\lambda} & \Phi^{\nu}{}_{\beta\mu} \end{vmatrix} = \begin{vmatrix} \delta^{\kappa}_{\nu}\chi_{\lambda}+\delta^{\kappa}_{\lambda}\chi_{\nu}-g^{\kappa\rho}g_{\nu\lambda}\chi_{\rho} & \delta^{\kappa}_{\nu}\chi_{\mu}+\delta^{\kappa}_{\mu}\chi_{\nu}-g^{\kappa\rho}g_{\nu\mu}\chi_{\rho} \\ \delta^{\nu}_{\beta}\chi_{\lambda}+\delta^{\nu}_{\lambda}\chi_{\beta}-g^{\nu\sigma}g_{\beta\lambda}\chi_{\sigma} & \delta^{\nu}_{\beta}\chi_{\mu}+\delta^{\nu}_{\mu}\chi_{\beta}-g^{\nu\sigma}g_{\beta\mu}\chi_{\sigma} \end{vmatrix}$$

$$= \begin{vmatrix} \delta^{\kappa}_{\lambda}\chi_{\nu} & \delta^{\kappa}_{\mu}\chi_{\nu} \\ \delta^{\nu}_{\lambda}\chi_{\beta} & \delta^{\nu}_{\mu}\chi_{\beta} \end{vmatrix} - \begin{vmatrix} \delta^{\kappa}_{\lambda}\chi_{\nu} & g^{\kappa\rho}g_{\nu\mu}\chi_{\rho} \\ \delta^{\nu}_{\lambda}\chi_{\beta} & g^{\nu\sigma}g_{\beta\mu}\chi_{\sigma} \end{vmatrix} - \begin{vmatrix} g^{\kappa\rho}g_{\nu\lambda}\chi_{\rho} & \delta^{\kappa}_{\mu}\chi_{\nu} \\ g^{\nu\sigma}g_{\beta\mu}\chi_{\sigma} & \delta^{\kappa}_{\lambda}\chi_{\beta} \end{vmatrix} + \begin{vmatrix} g^{\kappa\rho}g_{\nu\lambda}\chi_{\rho} & g^{\kappa\rho}g_{\nu\mu}\chi_{\rho} \\ g^{\nu\sigma}g_{\beta\lambda}\chi_{\sigma} & g^{\nu\sigma}g_{\beta\mu}\chi_{\sigma} \end{vmatrix}.$$

If we adopt the symbol $\Delta_1(\chi, \chi)$ defined in (5.4–10) we find

$$\begin{aligned} \overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu} &= R^{\kappa}{}_{\beta\lambda\mu}+\delta^{\kappa}_{\lambda}(\chi_{\beta\mu}-\chi_{\beta}\chi_{\mu})+\delta^{\kappa}_{\mu}(\chi_{\beta\lambda}-\chi_{\beta}\chi_{\lambda})+g^{\kappa\nu}g_{\beta\lambda}(\chi_{\nu\mu}-\chi_{\nu}\chi_{\mu}) \\ &= g^{\kappa\nu}g_{\beta\mu}(\chi_{\nu\lambda}-\chi_{\nu}\chi_{\lambda})+(\delta^{\kappa}_{\mu}g_{\beta\lambda}-\delta^{\kappa}_{\lambda}g_{\beta\mu})\Delta_1(\chi, \chi) \\ &= R^{\kappa}{}_{\beta\lambda\mu}-\delta^{\kappa}_{\lambda}(\chi_{\beta\mu}-\chi_{\beta}\chi_{\mu}+\tfrac{1}{2}g_{\beta\mu}\Delta_1(\chi, \chi))+ \\ &\qquad -\delta^{\kappa}_{\mu}(\chi_{\beta\lambda}-\chi_{\beta}\chi_{\lambda}+\tfrac{1}{2}g_{\beta\lambda}\Delta_1(\chi, \chi))+ \\ &+g^{\kappa\nu}g_{\beta\lambda}(\chi_{\nu\mu}-\chi_{\nu}\chi_{\mu}+\tfrac{1}{2}g_{\nu\mu}\Delta_1(\chi, \chi))-g^{\kappa\nu}g_{\beta\mu}(\chi_{\nu\lambda}-\chi_{\nu}\chi_{\lambda}+\tfrac{1}{2}g_{\nu\lambda}\Delta_1(\chi, \chi)). \end{aligned} \tag{9.6–5}$$

Introducing the abbreviation

$$\omega_{\lambda\mu} = \chi_{\lambda\mu}-\chi_{\lambda}\chi_{\mu}+\tfrac{1}{2}g_{\lambda\mu}\Delta_1(\chi, \chi) = \omega_{\mu\lambda}, \tag{9.6–6}$$

the equation (9.6–5) takes the form

$$\overset{*}{R}{}^{\kappa}{}_{\beta\lambda\mu} = R^{\kappa}{}_{\beta\lambda\mu} - \begin{vmatrix} \delta^{\kappa}_{\lambda} & \delta^{\kappa}_{\mu} \\ \omega_{\beta\lambda} & \omega_{\beta\mu} \end{vmatrix} - \begin{vmatrix} g^{\kappa\nu}\omega_{\lambda\nu} & g^{\kappa\nu}\omega_{\mu\nu} \\ g_{\beta\lambda} & g_{\beta\mu} \end{vmatrix}. \tag{9.6–7}$$

Contracting on κ and μ yields

$$\overset{*}{R}_{\beta\lambda} = R_{\beta\lambda}+(n-2)\omega_{\beta\lambda}+g_{\beta\lambda}\omega, \tag{9.6–8}$$

with

$$\omega = g^{\mu\nu}\omega_{\mu\nu}.$$

Transvecting both members of (9.6–8) by $\overset{*}{g}{}^{\beta\lambda}e^{2\chi} = g^{\beta\lambda}$ leads to

$$e^{2\chi}\overset{*}{R} = R+2(n-1)\omega \tag{9.6–9}$$

whence

$$\overset{*}{g}_{\lambda\mu}\overset{*}{R} = g_{\lambda\mu}R+2(n-1)g_{\lambda\mu}\omega. \tag{9.6–10}$$

Eliminating ω from (9.6–8) and (9.6–10) we get

$$2(n-1)\overset{*}{R}_{\beta\lambda}-\overset{*}{g}_{\beta\lambda}\overset{*}{R} = 2(n-1)R_{\beta\lambda}-g_{\beta\lambda}R+2(n-1)(n-2)\omega_{\beta\lambda}.$$

This may be written as

$$(n-2)\omega_{\beta\lambda} = \overset{*}{L}_{\beta\lambda}-L_{\beta\lambda}, \tag{9.6–11}$$

with

$$\boxed{L_{\lambda\mu} = R_{\lambda\mu}-\frac{1}{2(n-1)}g_{\lambda\mu}R.} \tag{9.6–12}$$

When we substitute the value of $\omega_{\lambda\mu}$ from (9.6–11) into (9.6–7), assuming thereby that $n > 2$, we find after an easy computation

$$\overset{*}{C}{}^{\kappa}{}_{\beta\lambda\mu} = C^{\kappa}{}_{\beta\lambda\mu}, \tag{9.6–13}$$

where

$$\boxed{C_{\alpha\beta\lambda\mu} = g_{\alpha\kappa}C^{\kappa}{}_{\beta\lambda\mu} = R_{\alpha\beta\lambda\mu}+\frac{1}{n-2}\left(\begin{vmatrix} g_{\alpha\lambda} & g_{\alpha\mu} \\ L_{\beta\lambda} & L_{\beta\mu}\end{vmatrix}+\begin{vmatrix} L_{\alpha\lambda} & L_{\alpha\mu} \\ g_{\beta\lambda} & g_{\beta\mu}\end{vmatrix}\right).} \tag{9.6–14}$$

This tensor is known as *the conformal curvature tensor* and was also discovered by H. Weyl. Summing up we have:

The mixed components of the conformal curvature tensor are the same for manifolds in conformal correspondence.

By the same arguments as those used in section 9.5.1 we may replace the Christoffel symbols involved in the expression for $C^{\kappa}{}_{\beta\lambda\mu}$ by the conformal Thomas symbols.

If the manifold is of constant Riemannian curvature then, according to (9.3–21) and (9.3–22),

$$L_{\lambda\mu} = \tfrac{1}{2}(2-n)Kg_{\lambda\mu}. \tag{9.6–15}$$

Inserting this into (9.6–14) we find

The conformal curvature tensor vanishes for manifolds with constant Riemannian curvature.

We complete this theorem by the assertion:

The conformal curvature tensor is zero for all three dimensional manifolds.

According to (9.2–7) the number of independent components of the Riemannian tensor of a three dimensional manifold is six. This is also the number of components $L_{\lambda\mu}$. We may consider them as unknowns in the system of equations

$$-R_{\alpha\beta\lambda\mu} = \begin{vmatrix} g_{\alpha\lambda} & g_{\alpha\mu} \\ L_{\beta\lambda} & L_{\beta\mu} \end{vmatrix} + \begin{vmatrix} L_{\alpha\lambda} & L_{\alpha\mu} \\ g_{\beta\lambda} & g_{\beta\mu} \end{vmatrix}. \tag{9.6–16}$$

By writing out these equations in full we obtain a system of six linear equations for the unknowns $L_{\lambda\mu}$. The determinant of this system is not identically zero as may be seen by inspection when we take the particular values $g_{11} = g_{22} = g_{33} = 1$, while the remaining g's are zero. Hence the system (9.6–16) is consistent and we can also find a solution in the general case. It follows that

$$-R_{\beta\lambda} = (1-3)L_{\beta\lambda}+L_{\beta\lambda}-g_{\beta\lambda}L = -L_{\beta\lambda}-g_{\beta\lambda}L$$

with

$$L = L_{\lambda\mu}g^{\lambda\mu}.$$

Hence

$$R = L+3L = 4L$$

and

$$L_{\beta\lambda} = R_{\beta\lambda}-\tfrac{1}{4}g_{\beta\lambda}R.$$

These are exactly the expressions (9.6–12) for $n = 3$.

9.6.2 – Einstein manifolds with vanishing conformal curvature tensor

We wish to establish the following theorem:

If the conformal curvature tensor of an Einstein manifold is zero, then the manifold is of constant Riemannian curvature.

The assertion is trivial when $n = 3$. In fact, a three dimensional Einstein manifold is always of constant Riemannian curvature as we pointed out in section 9.3.5. In the previous section we showed that the conformal curvature tensor of a three dimensional manifold is always zero.

Suppose that $n > 3$. By hypothesis we have

$$(2-n)R_{\alpha\beta\lambda\mu} = \begin{vmatrix} L_{\alpha\lambda} & L_{\alpha\mu} \\ g_{\beta\lambda} & g_{\beta\mu} \end{vmatrix} + \begin{vmatrix} g_{\alpha\lambda} & g_{\alpha\mu} \\ L_{\beta\lambda} & L_{\beta\mu} \end{vmatrix} \tag{9.6–17}$$

and

$$R_{\lambda\mu} = \frac{R}{n} g_{\lambda\mu}. \tag{9.6–18}$$

Hence

$$L_{\lambda\mu} = \frac{n-2}{2n(n-1)} Rg_{\lambda\mu} = \frac{n-2}{2n(n-1)} Rg_{\lambda\mu} = \tfrac{1}{2}(2-n)Kg_{\lambda\mu}.$$

Inserting this into (9.6–17) the truth of the statement follows easily.

CHAPTER 10

THEORY OF INTEGRABILITY

In this concluding chapter we wish to discuss some geometric problems which can be solved by applying the theory of integrability of partial differential equations of the first order. We shall not state explicitly the minimum conditions under which the results are valid, for they may be found in text-books on differential equations. For our purposes it is sufficient to suppose that the functions under consideration have continuous derivatives up to the order entering in the treatment throughout a certain region of the independent variables.

10.1 – The conditions of integrability

10.1.1 – COMPLETE SYSTEMS

Let us consider a system of partial differential equations

$$\partial_\mu x^\kappa = F^\kappa{}_\mu(x^1, \ldots, x^m;\ q^1, \ldots, q^n), \tag{10.1–1}$$

where κ runs through the symbols $1, \ldots, m$ and μ through the symbols $1, \ldots, n$.

Suppose this system has a set of solutions

$$x^\kappa(q^1, \ldots, q^n) \tag{10.1–2}$$

Inserting these functions into (10.1–1) and differentiating with respect to q^λ we get

$$\partial_{\lambda\mu} x^\kappa = \frac{\partial F^\kappa{}_\mu}{\partial x^\nu}\,\partial_\lambda x^\nu + \partial_\lambda F^\kappa{}_\mu = \frac{\partial F^\kappa{}_\mu}{\partial x^\nu}\,F^\nu{}_\lambda + \partial_\lambda F^\kappa{}_\mu, \tag{10.1–3}$$

where $\partial_\lambda F^\kappa{}_\mu$ denotes the partial derivative of $F^\kappa{}_\mu(x^1, \ldots, x^m;\ q^1, \ldots, q^n)$ with respect to q^λ, the $x^1, \ldots, x^m$ being considered as independent variables.

Since the left-hand members of these equations are symmetric with respect to λ and μ we see that it is necessary that the functions on the right of (10.1–1) be such that

$$\boxed{\partial_\lambda F^\kappa{}_\mu + \frac{\partial F^\kappa{}_\mu}{\partial x^\nu}\,F^\nu{}_\lambda = \partial_\mu F^\kappa{}_\lambda + \frac{\partial F^\kappa{}_\lambda}{\partial x^\nu}\,F^\nu{}_\mu,} \tag{10.1–4}$$

when the variables $x^1, \ldots, x^m$ are replaced by the functions (10.1–2). In general there are more relations (10.1–4) than functions x^κ. In fact, the number of equations (10.1–4) is $\frac{1}{2}mn(n-1) > m$, whenever $n \geqq 3$. Hence the system (10.1–1) is integrable only in particular circumstances.

The equations (10.1–4) are called the *conditions of integrability* of the system (10.1–1). An important case arises when they are identities in the variables x and q. The system (10.1–1) is then said to be *a complete system* and in the next section we shall show that it is *completely integrable*, i.e., that there exist solutions involving m arbitrary constants.

10.1.2 – The integration of a complete system

We proceed by establishing the following fundamental theorem:

A complete system of partial differential equations (10.1–1) *admits one and only one set of solutions* $x^\kappa(q^1, \ldots, q^n)$ *such that for arbitrary values* a^λ *of* q^λ *(within a certain region) the functions* x^κ *reduce to arbitrary constants* b^κ.

The theorem is true when $n = 1$, for then it is an assertion about a system of ordinary differential equations. In this case there are no conditions (10.1–4) and, therefore, the system may be considered as a complete system.

Let us assume the validity of the theorem for $n-1$ independent variables. We consider the system

$$\partial_\mu x^\kappa = F^\kappa{}_\mu(x^1, \ldots, x^m; q^1, \ldots, q^{n-1}, a^n), \tag{10.1–5}$$

where μ runs through the symbols $1, \ldots, n-1$. By hypothesis there is a set of solutions

$$x^\kappa(q^1, \ldots, q^{n-1}, a^n) \tag{10.1–6}$$

depending on the parameter a^n and reducing to the prescribed constants b^κ for $q^1 = a^1, \ldots, q^{n-1} = a^{n-1}$.

Next we consider the system

$$\partial_n x^\kappa = F^\kappa{}_n(x^1, \ldots, x^m; q^1, \ldots, q^n) \tag{10.1–7}$$

as a system of ordinary differential equations with respect to the variable q^n, the variables $q^1, \ldots, q^{n-1}$ taking the part of parameters. On account of the existence theorem for ordinary differential equations, we can find a set of functions

$$x^\kappa(q^1, \ldots, q^n) \tag{10.1–8}$$

satisfying the system (10.1–7) and reducing to the functions (10.1–6) for $q^n = a^n$. They have, moreover, continuous partial derivatives with respect to all variables. It remains to be verified that they are also solutions of the whole system (10.1–1).

If $\mu \neq n$ we have

$$\partial_\mu x^\kappa\big|_{q^n=a^n} = F^\kappa{}_\mu(x^1, \ldots, x^m; q^1, \ldots, q^{n-1}, a^n).$$

Hence

$$\partial_\mu x^\kappa = F^\kappa{}_\mu(x^1, \ldots, x^m; q^1, \ldots, q^n) + G^\kappa{}_\mu(q^1, \ldots, q^n). \tag{10.1–9}$$

Here $G^\kappa{}_\mu(q^1, \ldots, q^n)$ are functions which vanish identically when $q^n = a^n$. We shall see presently that they vanish identically when q^n is an arbitrary variable. To this end we perform the differentiation with respect to q^n:

$$\partial_n G^\kappa{}_\mu = \partial_{n\mu} x^\kappa - \frac{\partial F^\kappa{}_\mu}{\partial x^\nu} F^\nu{}_n - \partial_n F^\kappa{}_\mu, \tag{10.1–10}$$

where we have taken account of the fact that the x^κ are solutions of (10.1–7). In view of (10.1–9) we have

$$\begin{aligned}\partial_{n\mu} x^\kappa = \partial_{\mu n} x^\kappa &= \partial_\mu F^\kappa{}_n + \frac{\partial F^\kappa{}_n}{\partial x^\nu} \partial_\mu x^\nu \\ &= \partial_\mu F^\kappa{}_n + \frac{\partial F^\kappa{}_n}{\partial x^\nu} (F^\nu{}_\mu + G^\nu{}_\mu).\end{aligned}$$

By virtue of the conditions of integrability (10.1–4), these being identities, we may conclude that (10.1–10) reduces to

$$\partial_n G^\kappa{}_\mu = \frac{\partial F^\kappa{}_n}{\partial x^\nu} G^\nu{}_\mu. \tag{10.1–11}$$

Thus we see that the functions $G^\kappa{}_\mu$, considered as functions of q^n, are solutions of a system of ordinary differential equations and these solutions reduce to zero for $q^n = a^n$. Since a solution is uniquely determined by its initial values we may infer that the $G^\nu{}_\mu$ are identically zero and this concludes the proof of the theorem.

10.1.3 – Condition for a gradient

A very simple illustration of the theorem of the previous section is concerned with the problem of finding conditions under which a covariant vector field coincides with a gradient. If the $v_\mu(q^1, \ldots, q^n)$ are the components of the field we have to consider the system of partial differential equations

$$\partial_\mu x = v_\mu. \tag{10.1–12}$$

The conditions of integrability are

$$\partial_\lambda v_\mu = \partial_\mu v_\lambda, \tag{10.1–13}$$

stating that the rotation of $\boldsymbol{v}$ vanishes. Hence:

A covariant vector field is a gradient if and only if its rotation vanishes identically.

10.1.4 – TENSOR EQUATIONS

In many geometrical equations we encounter equations of the type

$$\nabla_\mu x^{\kappa_1 \dots \kappa_p}{}_{\beta_1 \dots \beta_q} = f_\mu{}^{\kappa_1 \dots \kappa_p}{}_{\beta_1 \dots \beta_q}, \tag{10.1–14}$$

where the $x^{\kappa_1 \dots \kappa_p}{}_{\beta_1 \dots \beta_q}$ are the components of a tensor and unknown functions of the variables q^λ. It is understood that these equations have a meaning on a given n-dimensional manifold and that ∇_μ symbolizes the covariant derivative. The expressions on the right, being also the components of a tensor, depend on the q's as well as on the x's.

First we consider a particular case:

$$\nabla_\mu x_\beta = f_{\mu\beta}. \tag{10.1–15}$$

This system is essentially a system of partial differential equations, viz.,

$$\partial_\mu x_\beta = f_{\mu\beta} + \Gamma^\nu{}_{\beta\mu} x_\nu. \tag{10.1–16}$$

The conditions of integrability arise from

$$\partial_\lambda f_{\mu\beta} + \frac{\partial f_{\mu\beta}}{\partial x_\nu} \partial_\lambda x_\nu + \partial_\lambda \Gamma^\nu{}_{\beta\mu} x_\nu + \Gamma^\nu{}_{\beta\mu} \partial_\lambda x_\nu$$

$$= \partial_\mu f_{\lambda\beta} + \frac{\partial f_{\lambda\beta}}{\partial x_\nu} \partial_\mu x_\nu + \partial_\mu \Gamma^\nu{}_{\beta\lambda} x_\nu + \Gamma^\nu{}_{\beta\lambda} \partial_\mu x_\nu,$$

where the derivatives of the x's must be replaced by their expressions from (10.1–16). Performing the substitutions and taking account of (9.4–2) a straightforward computation yields

$$-R^\nu{}_{\beta\lambda\mu} x_\nu = \nabla_\lambda f_{\mu\beta} - \nabla_\mu f_{\lambda\beta}, \tag{10.1–17}$$

where $\nabla_\lambda f_{\mu\beta}$ stands for

$$\partial_\lambda f_{\mu\beta} + \frac{\partial f_{\mu\beta}}{\partial x_\nu} \partial_\lambda x_\nu - \Gamma^\nu{}_{\mu\lambda} f_{\nu\beta} - \Gamma^\nu{}_{\beta\lambda} f_{\mu\nu}$$

and $\partial_\lambda x_\nu$ for the expression on the right of (10.1–16) (with appropriate subscripts, of course). A similar expression stands for $\nabla_\mu f_{\lambda\beta}$.

With reference to (9.2–28) we may state that *the conditions of integrability* (10.1–17) *have been obtained by formal covariant differentiation of* (10.1–15).

Proceeding along the same lines we find that the conditions of integrability of the system

$$\nabla_\mu x^\kappa = f_\mu{}^\kappa \tag{10.1–18}$$

may be written as

$$R^\kappa{}_{\nu\lambda\mu} x^\nu = \nabla_\lambda f_\mu{}^\kappa - \nabla_\mu f_\lambda{}^\kappa. \tag{10.1–19}$$

The conditions of integrability of the system (10.1–14) will be found when we take account of the remark at the end of section 9.2.5. The differentiated terms of $\nabla_\lambda f_\mu{}^{\kappa_1 \ldots \kappa_p}{}_{\beta_1 \ldots \beta_q}$ are

$$\partial_\lambda f_\mu{}^{\kappa_1 \ldots \kappa_p}{}_{\beta_1 \ldots \beta_q} + \frac{\partial f_\mu{}^{\kappa_1 \ldots \kappa_p}{}_{\beta_1 \ldots \beta_q}}{\partial x^{\kappa_1 \ldots \kappa_p}{}_{\beta_1 \ldots \beta_q}} \partial_\lambda x^{\nu_1 \ldots \nu_p}{}_{\rho_1 \ldots \rho_q},$$

where the $\partial_\lambda x^{\nu_1 \ldots \nu_p}{}_{\rho_1 \ldots \rho_q}$ are obtained from (10.1–14).

10.1.5 – Manifolds isometrically equivalent to a flat

It is an easy matter to show that the components of the Riemannian tensor of a flat are identically zero. In fact, a flat may be represented by the equations (9.6–1). As a consequence the first derivatives of the coordinate vector are constant and, consequently, the second derivatives vanish.

It is our aim to establish the converse theorem, that is to say:

A manifold such that the components of its Riemannian tensor are identically zero is isometrically equivalent to a flat.

First we observe that the system

$$\nabla_\mu x_\beta = 0 \tag{10.1–20}$$

where x_β represents the components of a vector field on the manifold, is completely integrable. Indeed, the conditions of integrability are

$$R^\nu{}_{\beta\lambda\mu} x_\nu = 0 \tag{10.1–21}$$

and, by hypothesis, these are identically fulfilled. Hence, by taking suitable initial values, we can find a set of solutions

$$q^1_\beta, \ldots, q^n_\beta, \qquad \beta = 1, \ldots, n, \tag{10.1–22}$$

such that

$$\det [q^k_\beta] \neq 0, \qquad k = 1, \ldots, n, \quad \beta = 1, \ldots, n, \tag{10.1–23}$$

at the point determined by initial values of the independent variables q^κ.

We shall write

$$\partial_\mu q^k_\beta = \Gamma^\kappa{}_{\lambda\mu} q^k_\kappa \tag{10.1–24}$$

rather than

$$\nabla_\mu q^k_\beta = 0.$$

Next we consider the system of partial differential equations

$$\partial_\mu q^k = q^k_\mu, \tag{10.1–25}$$

the q^k, q^k_μ still being functions of the q^κ. The conditions of integrability are

$$\partial_\lambda q^k_\mu = \partial_\mu q^k_\lambda \tag{10.1–26}$$

and these have been identically fulfilled because of (10.1–24). Hence we can find a set of functions q^k such that the jacobian, being the determinant (10.1–23), does not vanish. We may take them as a new set of parameters. In view of the law of transformation (7.4–3) we have

$$\Gamma^\kappa{}_{\lambda\mu}\partial_\kappa q^k = \Gamma^k{}_{lm}\partial_\lambda q^l \partial_\mu q^m + \partial_{\lambda\mu} q^k.$$

From (10.1–25), (10.1–24) and (10.1–23) we may infer that

$$\Gamma^k{}_{lm} = 0, \qquad k, l, m = 1, \ldots, n,$$

throughout the manifold.

If g_{lm} are the components of the metric tensor with respect to the parameter system (k) we find from Ricci's theorem (9.1–44)

$$\partial_k g_{lm} = \nabla_k g_{lm} = 0.$$

Hence the components are constant throughout the manifold and it is, therefore, isometrically equivalent to a flat.

10.2 – Geodesic and conformal mapping

10.2.1 – Geodesically equivalent manifolds

Let us assume that a manifold $\overset{*}{x}(q^\kappa)$ is geodesically equivalent to a given manifold $x(q^\kappa)$. We wish to investigate the metric tensor $\overset{*}{g}_{\lambda\mu}$ of the manifold $\overset{*}{x}(q^\kappa)$.

Ricci's theorem of section 9.1.7 states that the covariant derivative of the metric tensor $\overset{*}{g}_{\lambda\mu}$ is zero on the manifold $\overset{*}{x}$. This can be expressed by the equation

$$\partial_\kappa \overset{*}{g}_{\lambda\mu} = \overset{*}{\Gamma}{}^\nu{}_{\lambda\kappa}\overset{*}{g}_{\nu\mu} + \overset{*}{\Gamma}{}^\nu{}_{\mu\kappa}\overset{*}{g}_{\lambda\nu}. \tag{10.2–1}$$

Next we refer to the relation (7.6–9). Accordingly (10.2–1) may be cast into the form

$$\nabla_\kappa \overset{*}{g}_{\lambda\mu} = 2\varphi_\kappa \overset{*}{g}_{\lambda\mu} + \varphi_\lambda \overset{*}{g}_{\mu\kappa} + \varphi_\mu \overset{*}{g}_{\kappa\lambda}, \tag{10.2–2}$$

where the covariant differentiation has been defined by means of the Christoffel symbols on the manifold x. The equation (10.2–2) expresses that the components of the metric tensor of the manifold $\overset{*}{x}$ satisfy a system of partial differential equations.

Conversely, let us assume that we can find a vector φ_λ and a manifold $\overset{*}{x}$ whose metric tensor satisfies (10.2–2), where it is understood that this second manifold is referred to the parameters of the manifold x. Observing that

$$\tfrac{1}{2}(\nabla_\lambda \overset{*}{g}_{\mu\kappa}+\nabla_\mu \overset{*}{g}_{\kappa\lambda}-\nabla_\kappa \overset{*}{g}_{\lambda\mu}) = \overset{*}{\Gamma}_{\kappa\lambda\mu}-\Gamma^\nu{}_{\lambda\mu}\overset{*}{g}_{\nu\kappa}$$

we find from (10.2–2) that

$$\overset{*}{\Gamma}_{\kappa\lambda\mu}=\Gamma^\nu{}_{\lambda\mu}\overset{*}{g}_{\nu\kappa}+\varphi_\lambda \overset{*}{g}_{\mu\kappa}+\varphi_\mu \overset{*}{g}_{\kappa\nu\lambda}. \tag{10.2–3}$$

Transvection by $\overset{*}{g}{}^{\kappa\nu}$ yields the equations (7.6–9) and it is easy to show that the φ_λ are the components of a gradient. Hence

A manifold $\overset{*}{x}(q^\kappa)$ *is geodesically equivalent to a manifold* $x(q^\kappa)$ *if and only if the components* $\overset{*}{g}_{\lambda\mu}$ *of its metric tensor are solutions of a system of partial differential equations* (10.2–2), *where the* φ_κ *are the components of a suitably chosen gradient.*

We proceed by writing down the conditions of integrability of the system (10.2–2). They arise from

$$\nabla_\lambda\nabla_\mu \overset{*}{g}_{\beta\gamma}-\nabla_\mu\nabla_\lambda \overset{*}{g}_{\beta\gamma} = \nabla_\lambda(2\varphi_\mu \overset{*}{g}_{\beta\gamma}+\varphi_\beta \overset{*}{g}_{\gamma\mu}+\varphi_\gamma \overset{*}{g}_{\mu\beta})+$$
$$-\nabla_\mu(2\varphi_\lambda \overset{*}{g}_{\beta\gamma}+\varphi_\beta \overset{*}{g}_{\gamma\lambda}+\varphi_\gamma \overset{*}{g}_{\lambda\beta}),$$

where the expression on the left must be replaced by

$$-R^\nu{}_{\beta\lambda\mu}\overset{*}{g}_{\nu\gamma}-R^\nu{}_{\gamma\lambda\mu}\overset{*}{g}_{\nu\beta}$$

and the covariant derivatives of the $\overset{*}{g}$'s on the right by their expressions (10.2–2). From

$$\nabla_\lambda\nabla_\mu \overset{*}{g}_{\beta\gamma} = 2\varphi_{\lambda\mu}\overset{*}{g}_{\beta\gamma}+\varphi_{\beta\lambda}\overset{*}{g}_{\gamma\mu}+\varphi_{\gamma\lambda}\overset{*}{g}_{\mu\beta}+2\varphi_\mu(2\varphi_\lambda \overset{*}{g}_{\beta\gamma}+\varphi_\beta \overset{*}{g}_{\gamma\lambda}+\varphi_\gamma \overset{*}{g}_{\lambda\beta})$$
$$+\varphi_\beta(2\varphi_\lambda \overset{*}{g}_{\gamma\mu}+\varphi_\gamma \overset{*}{g}_{\mu\lambda}+\varphi_\mu \overset{*}{g}_{\lambda\gamma})+\varphi_\gamma(2\varphi_\lambda \overset{*}{g}_{\beta\mu}+\varphi_\mu \overset{*}{g}_{\beta\lambda}+\varphi_\beta \overset{*}{g}_{\lambda\mu})$$

we get by introducing the abbreviation (9.5–2):

$$R^\nu{}_{\beta\lambda\mu}\overset{*}{g}_{\nu\gamma}+\psi_{\beta\lambda}\overset{*}{g}_{\gamma\mu}-\psi_{\beta\mu}\overset{*}{g}_{\gamma\lambda}+R^\nu{}_{\gamma\lambda\mu}\overset{*}{g}_{\nu\beta}+\psi_{\gamma\lambda}\overset{*}{g}_{\beta\mu}-\psi_{\gamma\mu}\overset{*}{g}_{\beta\lambda} = 0. \tag{10.2–4}$$

These are the required conditions of integrability.

10.2.2 – Geodesically flat manifolds

In this section we shall complete the discussion of section 9.5.2.

A manifold is said to be *geodesically flat* when it is geodesically equivalent to a flat (or, what amounts to the same thing, when it is geodesically equivalent to a manifold which is itself isometrically equivalent to a flat).

Suppose that the manifold $\overset{*}{x}$ is isometrically flat. Under the assumption that the manifold x is geodesically equivalent to $\overset{*}{x}$ the relation (9.5–3) yields

$$R^{\kappa}{}_{\beta\lambda\mu} = \delta^{\kappa}_{\lambda}\psi_{\beta\mu} - \delta^{\kappa}_{\mu}\psi_{\beta\lambda} \tag{10.2–5}$$

and (9.5–2) reads

$$\nabla_{\mu}\varphi_{\beta} = \psi_{\mu\beta} + \varphi_{\mu}\varphi_{\beta}. \tag{10.2–6}$$

Assume, conversely, that the Riemannian tensor of the manifold x takes the form (10.2–5). Contraction with respect to κ and μ yields

$$R_{\beta\lambda} = (1-n)\psi_{\beta\lambda} \tag{10.2–7}$$

and it follows that the $\psi_{\beta\lambda}$ are the components of a symmetric tensor. According to section 10.1.3 every vector φ_{β} satisfying (10.2–6) is a gradient.

The assumption stated above is equivalent to the statement that the mixed components (9.5–6) of Weyl's tensor vanish. Hence, according to the theorem of section 9.3.6, the manifold is of constant Riemannian curvature whenever $n > 2$. Thus the assumption (10.2–5) asserts that the Riemannian tensor takes the form

$$R^{\kappa}{}_{\beta\lambda\mu} = K(\delta^{\kappa}_{\lambda}g_{\beta\mu} - \delta^{\mu}_{\kappa}g_{\beta\lambda}), \tag{10.2–8}$$

where K is constant, provided that $n > 2$.

Suppose next that the equations (10.2–6) have a solution. If the functions φ_{κ} are inserted into the equations (10.2–2) then it is an easy matter to verify that they also have a solution, for from (10.2–5) follows that the relations (10.2–4) are identities. Hence the manifold x is geodesically equivalent to a manifold $\overset{*}{x}$ whose metric tensor has the components $\overset{*}{g}_{\lambda\mu}$. By virtue of (9.5–3) the components of the Riemannian tensor of the manifold $\overset{*}{x}$ are identically zero and, consequently, the manifold x appears to be geodesically flat.

It remains to discuss the equations (10.2–6). According to (10.1–17) the conditions of integrability are

$$-R^{\nu}{}_{\beta\lambda\mu}\varphi_{\nu} = \nabla_{\lambda}\psi_{\mu\beta} - \nabla_{\mu}\psi_{\lambda\beta} + \nabla_{\lambda}(\varphi_{\mu}\varphi_{\beta}) - \nabla_{\mu}(\varphi_{\lambda}\varphi_{\beta}). \tag{10.2–9}$$

Taking account of (10.2–5) and (10.2–6) a straightforward computation yields

$$\nabla_\lambda \psi_{\mu\beta} - \nabla_\mu \psi_{\lambda\beta} = 0. \tag{10.2–10}$$

It is interesting to notice that *these relations are a consequence of* (10.2–5) *in the case that* $n > 2$. In fact, applying Bianchi's identity (9.2–13) to (10.2–5) we get

$$\delta^\kappa_\lambda \nabla_\nu \psi_{\beta\mu} - \delta^\kappa_\mu \nabla_\nu \psi_{\beta\lambda} + \delta^\kappa_\mu \nabla_\lambda \psi_{\beta\nu} - \delta^\kappa_\nu \nabla_\lambda \psi_{\beta\mu} + \delta^\kappa_\nu \nabla_\mu \psi_{\beta\lambda} - \delta^\kappa_\lambda \nabla_\mu \psi_{\beta\nu} = 0.$$

Contraction with respect to κ and ν yields

$$(n-2)(\nabla_\mu \psi_{\beta\lambda} - \nabla_\lambda \psi_{\beta\mu}) = 0$$

and this proves the assertion.

From (10.2–7) and (10.2–8) we deduce

$$\psi_{\lambda\mu} = K g_{\lambda\mu}. \tag{10.2–11}$$

Hence (10.2–10) is equivalent to

$$g_{\mu\beta} \partial_\lambda K - g_{\lambda\beta} \partial_\mu K = 0,$$

whence

$$(n-1)\partial_\lambda K = 0.$$

The meaning of this is that the conditions (10.2–10) are identically satisfied whenever K is constant (including the case $n = 2$). Thus we proved:

Any manifold of constant Riemannian curvature is geodesically flat.

A direct consequence is:

Any two manifolds of constant Riemannian curvature are geodesically equivalent.

This result completes Beltrami's theorem of section 9.5.2.

Transvecting both members of (10.2–6) by $g^{\mu\beta}$, taking account of (10.2–11) we see that the scalar invariant φ whose derivatives appear in (10.2–2) satisfies the equation

$$\Delta\varphi - \Delta_1(\varphi, \varphi) = nK, \tag{10.2–12}$$

where $\Delta\varphi$ is the laplacian and $\Delta_1(\varphi, \varphi)$ is Beltrami's differential parameter (6.4–10).

10.2.3 – Conformally flat manifolds

A manifold which can be mapped conformally on a flat is said to be *conformally flat*. In section 9.6.1 we proved that the conformal curvature of such a manifold vanishes identically. It will be our task to find a sufficient condition for a manifold to be conformally flat.

It appears from (9.6–7) that there must exist a tensor $\omega_{\lambda\mu}$ such that

$$R_{\alpha\beta\lambda\mu} = \begin{vmatrix} g_{\alpha\lambda} & g_{\alpha\mu} \\ \omega_{\beta\lambda} & \omega_{\beta\mu} \end{vmatrix} + \begin{vmatrix} \omega_{\alpha\lambda} & \omega_{\alpha\mu} \\ g_{\beta\lambda} & g_{\beta\mu} \end{vmatrix}, \tag{10.2–13}$$

whence

$$R_{\beta\lambda} = -(n-2)\omega_{\beta\lambda} - g_{\beta\lambda}\,\omega \tag{10.2–14}$$

with

$$\omega = g^{\alpha\mu}\,\omega_{\alpha\mu}.$$

According to (10.2–14) the tensor $\omega_{\lambda\mu}$ is necessarily symmetric. In addition, according to (9.6–6), there must exist a scalar invariant χ such that

$$\nabla_\mu \chi_\beta = \omega_{\beta\mu} + \chi_\beta \chi_\mu - \tfrac{1}{2} g_{\beta\mu} \Delta_1(\chi, \chi), \tag{10.2–15}$$

where χ_κ stands for $\partial_\kappa \chi$.

The conditions of integrability of this system of partial differential equations arise from

$$\begin{aligned} -R^\nu{}_{\beta\lambda\mu}\chi_\nu &= \nabla_\lambda \omega_{\beta\mu} - \nabla_\mu \omega_{\beta\lambda} + \\ &+ \nabla_\lambda(\chi_\beta\chi_\mu) - \nabla_\mu(\chi_\beta\chi_\lambda) + \tfrac{1}{2} g_{\beta\lambda} \nabla_\lambda \Delta_1(\chi, \chi) - \tfrac{1}{2} g_{\beta\mu} \nabla_\lambda \Delta_1(\chi, \chi). \end{aligned} \tag{10.2–16}$$

First we observe that (10.2–14) implies

$$R^\nu{}_{\beta\lambda\mu}\chi_\nu = \chi_\lambda \omega_{\beta\mu} - \chi_\mu \omega_{\beta\lambda} + g^{\alpha\nu}\chi_\nu(\omega_{\alpha\lambda} g_{\beta\mu} - \omega_{\alpha\mu} g_{\beta\lambda}).$$

In view of (10.2–15) we have

$$\begin{aligned} \nabla_\lambda(\chi_\beta\chi_\mu) - \nabla_\mu(\chi_\beta\chi_\lambda) &= \chi_\mu \nabla_\lambda \chi_\beta - \chi_\lambda \nabla_\mu \chi_\beta \\ &= \chi_\mu \omega_{\beta\lambda} - \chi_\lambda \omega_{\beta\mu} + \tfrac{1}{2}(g_{\beta\mu}\chi_\lambda - g_{\beta\lambda}\chi_\mu)\Delta_1(\chi, \chi). \end{aligned}$$

Hence (10.2–16) reduces to

$$\begin{aligned} &\nabla_\lambda \omega_{\beta\mu} - \nabla_\mu \omega_{\beta\lambda} + g^{\alpha\nu}\chi_\nu(\omega_{\alpha\lambda} g_{\beta\mu} - \omega_{\alpha\mu} g_{\beta\lambda}) + \\ &+ \tfrac{1}{2} g_{\mu\beta}\chi_\lambda \Delta_1(\chi, \chi) - \tfrac{1}{2} g_{\lambda\beta}\chi_\mu \Delta_1(\chi, \chi) - \tfrac{1}{2} g_{\mu\beta} \nabla_\lambda \Delta_1(\chi, \chi) + \tfrac{1}{2} g_{\lambda\beta} \nabla_\mu \Delta_1(\chi, \chi). \end{aligned}$$

Finally we have

$$\begin{aligned} \tfrac{1}{2}\nabla_\mu \Delta_1(\chi, \chi) &= \tfrac{1}{2} g^{\alpha\nu} \nabla_\mu(\chi_\alpha \chi_\nu) = \tfrac{1}{2} g^{\alpha\nu}(\omega_{\alpha\mu} + \chi_\alpha\chi_\mu - \tfrac{1}{2} g_{\alpha\mu}\Delta_1(\chi, \chi))\chi_\nu + \\ &+ \tfrac{1}{2} g^{\alpha\nu}(\omega_{\nu\mu} + \chi_\nu\chi_\mu - \tfrac{1}{2} g_{\nu\mu}\Delta_1(\chi, \chi))\chi_\alpha = g^{\alpha\nu}\chi_\nu \omega_{\alpha\mu} + \tfrac{1}{2}\chi_\mu \Delta_1(\chi, \chi). \end{aligned}$$

Thus we may infer that the conditions of integrability of the system (10.2–15) take the simple form

$$\nabla_\lambda \omega_{\beta\mu} - \nabla_\mu \omega_{\beta\lambda} = 0. \tag{10.2–17}$$

Introducing the tensor $L_{\lambda\mu}$ defined in (9.6–12) we easily deduce from

(10.2–13) that

$$L_{\lambda\mu} = -(n-2)\omega_{\lambda\mu}. \tag{10.2–18}$$

Hence (10.2–17) is equivalent to

$$\nabla_\lambda L_{\beta\mu} - \nabla_\mu L_{\beta\lambda} = 0, \tag{10.2–19}$$

provided that $n > 2$. Inserting (10.2–18) into (10.2–13) we get

$$R_{\alpha\beta\lambda\mu} = -\frac{1}{n-2}\left(\begin{vmatrix} g_{\alpha\lambda} & g_{\alpha\mu} \\ L_{\beta\lambda} & L_{\beta\mu} \end{vmatrix} + \begin{vmatrix} L_{\alpha\lambda} & L_{\alpha\mu} \\ g_{\beta\lambda} & g_{\beta\mu} \end{vmatrix}\right). \tag{10.2–20}$$

Hence the assumption (10.2–13) is equivalent to the statement that the conformal curvature tensor vanishes.

Transvecting by $g^{\alpha\mu}$ we find

$$R_{\beta\lambda} = L_{\beta\lambda} + \frac{1}{n-2} L g_{\beta\lambda}, \tag{10.2–21}$$

where L stands for $L_{\lambda\mu}g^{\lambda\mu}$. Hence

$$R = L + \frac{n}{n-2}L = \frac{2n-2}{n-2}L. \tag{10.2–22}$$

As a consequence Einstein's tensor (9.2–16) appears as

$$G_{\lambda\mu} = L_{\lambda\mu} - L g_{\lambda\mu}. \tag{10.2–23}$$

Next we wish to show that *in the case that* $n > 3$ *the relations* (10.2–19) *follow from* (10.2–20). In fact, applying Bianchi's identity to (10.2–20) we get

$$\begin{aligned} &g_{\alpha\lambda}\nabla_\nu L_{\beta\mu} - g_{\alpha\mu}\nabla_\nu L_{\beta\lambda} + g_{\beta\mu}\nabla_\nu L_{\alpha\lambda} - g_{\beta\lambda}\nabla_\nu L_{\alpha\mu} \\ &+ g_{\alpha\mu}\nabla_\lambda L_{\beta\nu} - g_{\alpha\nu}\nabla_\lambda L_{\beta\mu} + g_{\beta\nu}\nabla_\lambda L_{\alpha\mu} - g_{\beta\mu}\nabla_\lambda L_{\alpha\nu} \\ &+ g_{\alpha\nu}\nabla_\mu L_{\beta\lambda} - g_{\alpha\lambda}\nabla_\mu L_{\beta\nu} + g_{\beta\lambda}\nabla_\mu L_{\alpha\nu} - g_{\beta\nu}\nabla_\mu L_{\alpha\lambda} = 0. \end{aligned}$$

Transvection by $g^{\alpha\nu}$ yields

$$\begin{aligned} &(n-2)(\nabla_\lambda L_{\beta\mu} - \nabla_\mu L_{\beta\lambda}) \\ &= g_{\beta\mu}(g^{\alpha\nu}\nabla_\nu L_{\alpha\lambda} - \nabla_\lambda L) - g_{\beta\lambda}(g^{\alpha\nu}\nabla_\nu L_{\alpha\mu} - \nabla_\mu L) \\ &= g_{\beta\mu}g^{\alpha\nu}\nabla_\nu G_{\alpha\lambda} - g_{\beta\lambda}g^{\alpha\nu}\nabla_\nu G_{\alpha\mu}. \end{aligned}$$

This last expression vanishes on account of (9.2–17).

Summing up we may state the following *Weyl-Schouten theorem*:

A manifold of dimension > 3 *is conformally flat if and only if the components of its conformal curvature tensor vanish identically. A three-dimensional manifold is conformally flat when the tensor* $L_{\lambda\mu}$ *satisfies the conditions* (10.2–19).

As we know, a surface is always conformally flat.

10.3 – Bonnet's problem

10.3.1 – Statement of the problem

An interesting application of the theory of integrability is concerned with the following problem: We are given two sets of functions

$$g_{\lambda\mu}(q^1, \ldots, q^n), \qquad g_{\lambda\mu} = g_{\mu\lambda}, \qquad \lambda, \mu = 1, \ldots, n, \tag{10.3–1}$$

and

$$h_{\lambda\mu}(q^1, \ldots, q^n), \qquad h_{\lambda\mu} = h_{\mu\lambda}, \qquad \lambda, \mu = 1, \ldots, n. \tag{10.3–2}$$

We ask whether there is a hypersurface

$$\boldsymbol{x}(q^1, \ldots, q^n) \tag{10.3–3}$$

in an $(n+1)$-dimensional space having (10.3–1) and (10.3–2) as fundamental tensors. This problem is known as *Bonnet's problem.*

We recall the equations of Gauss and Weingarten, viz.,

$$\partial_{\lambda\mu}\boldsymbol{x} = \partial_\kappa \boldsymbol{x} \Gamma^\kappa{}_{\lambda\mu} + \boldsymbol{n} h_{\lambda\mu} \tag{10.3–4}$$

and

$$\partial_\lambda \boldsymbol{n} = -\partial_\nu \boldsymbol{x} h_{\lambda\kappa} g^{\kappa\nu}. \tag{10.3–5}$$

These equations describe a relation between the fundamental tensors, the coordinate vector $\boldsymbol{x}$ and the normal vector $\boldsymbol{n}$.

Between the first and the second fundamental tensor exist, among other things, the relations of Gauss and Codazzi, viz.

$$\partial_\lambda h_{\kappa\mu} + \Gamma^\nu{}_{\kappa\mu} h_{\lambda\nu} = \partial_\mu h_{\kappa\lambda} + \Gamma^\nu{}_{\kappa\lambda} h_{\mu\nu} \tag{10.3–7}$$

and

$$R_{\alpha\beta\lambda\mu} = h_{\alpha\lambda} h_{\beta\mu} - h_{\alpha\mu} h_{\beta\lambda}, \tag{10.3–8}$$

where $R_{\alpha\beta\lambda\mu}$ is the expression (9.4–5).

If we make the additional assumption that

$$\det [g_{\lambda\mu}] > 0 \tag{10.3–9}$$

we can prove that the conditions (10.3–7) and (10.3–8) are sufficient in order that the sets (10.3–1) and (10.3–2) be realizable as the components of the first and the second fundamental tensor of a hypersurface.

Before starting our problem in a consistent way we wish to make some preliminary remarks.

First we observe that we can derive from the functions (10.3–1) a set of functions

$$g^{\lambda\mu}(q^1, \ldots, q^n) \tag{10.3–10}$$

such that

$$g_{\lambda\nu}g^{\nu\mu} = \delta^{\mu}_{\lambda} \tag{10.3–11}$$

identically. This is an algebraic problem, for we have to solve systems of linear equations with non-vanishing system determinants.

Next we introduce the three index symbols by means of their expressions (7.3–12) and (7.3–7). Again we introduce $R^{\kappa}{}_{\beta\lambda\mu}$ by means of (9.4–2) and $R_{\alpha\beta\lambda\mu}$ as $g_{\alpha\kappa}R^{\kappa}{}_{\beta\lambda\mu}$. It is now clear that the equations (10.3–7) and (10.3–8) have a meaning without reference to a hypersurface.

The computations in the subsequent sections are much facilitated by using the notion of covariant differentiation in a formal way by means of their expressions derived in section 9.1.4. A mere verification leads to equations such as

$$\nabla_{\lambda}\nabla_{\mu}x_{\beta}-\nabla_{\mu}\nabla_{\lambda}x_{\beta} = -R^{\nu}{}_{\beta\lambda\mu}x_{\nu}$$

if

$$\nabla_{\lambda}x_{\mu} = \partial_{\lambda}x_{\mu}-\Gamma^{\kappa}{}_{\lambda\mu}x_{\kappa}$$

and so on.

For the sake of illustration we wish to carry out the verification in this example. According to the definition we have

$$\begin{aligned}\nabla_{\lambda}\nabla_{\mu}x_{\beta} &= \partial_{\lambda}\nabla_{\mu}x_{\beta}-\Gamma^{\nu}{}_{\mu\lambda}\nabla_{\nu}x_{\beta}-\Gamma^{\nu}{}_{\beta\lambda}\nabla_{\mu}x_{\nu}\\ &= \partial_{\lambda\mu}x_{\beta}-\partial_{\lambda}\Gamma^{\nu}{}_{\beta\mu}x_{\nu}-\Gamma^{\nu}{}_{\beta\mu}\partial_{\lambda}x_{\mu}\\ &\quad -\Gamma^{\nu}{}_{\mu\lambda}\nabla_{\nu}x_{\beta}-\Gamma^{\nu}{}_{\beta\lambda}\partial_{\mu}x_{\nu}+\Gamma^{\nu}{}_{\beta\lambda}\Gamma^{\nu}{}_{\beta\mu}x_{\kappa}.\end{aligned}$$

Hence

$$\nabla_{\lambda}\nabla_{\mu}x_{\beta}-\nabla_{\mu}\nabla_{\lambda}x_{\beta} = -(\partial_{\lambda}\Gamma^{\kappa}{}_{\beta\mu}-\partial_{\mu}\Gamma^{\kappa}{}_{\beta\lambda}+\Gamma^{\kappa}{}_{\nu\mu}\Gamma^{\nu}{}_{\beta\lambda}-\Gamma^{\kappa}{}_{\nu\lambda}\Gamma^{\nu}{}_{\beta\mu})x_{\kappa}$$

and this proves the above assertion. Thus the whole apparatus of tensor calculus is at our disposal.

10.3.2 – The differential equations of Bonnet's problem

We proceed by showing that Bonnet's problem can be solved by means of a system of differential equations of the type (10.3–4) and (10.3–5).

Let $\boldsymbol{x}(q^{\kappa})$ and $\boldsymbol{y}(q^{\kappa})$, $\kappa = 1, \ldots, n$, denote two unknown vector functions which satisfy the system of differential equations

$$\partial_{\mu}\boldsymbol{x} = \boldsymbol{x}_{\mu}, \tag{10.3–12}$$

$$\partial_{\mu}\boldsymbol{x}_{\beta} = \boldsymbol{x}_{\kappa}\Gamma^{\kappa}{}_{\beta\mu}+\boldsymbol{y}h_{\mu\beta}, \tag{10.3–13}$$

$$\partial_{\mu}\boldsymbol{y} = -\boldsymbol{x}_{\kappa}h_{\mu\beta}g^{\beta\kappa}, \tag{10.3–14}$$

where the $\boldsymbol{x}_{\kappa}$, $\kappa = 1, \ldots, n$, are auxiliary functions, also unknown before, of course.

In a more concise form this system appears as

$$\nabla_\mu \boldsymbol{x} = \boldsymbol{x}_\mu, \tag{10.3–15}$$

$$\nabla_\mu \boldsymbol{x}_\beta = \boldsymbol{y} h_{\mu\beta}, \tag{10.3–16}$$

$$\nabla_\mu \boldsymbol{y} = -\boldsymbol{x}_\kappa h_{\mu\alpha} g^{\alpha\kappa}. \tag{10.3–17}$$

The fact that we are dealing with vector functions is of no importance. Each vector function represents a set of $n+1$ scalar functions, the components with respect to a fixed frame in the $(n+1)$-dimensional space. Hence the theory of integrability applies also to systems such as those under consideration.

Next we assume that the conditions (10.3–7) and (10.3–8) are identically fulfilled.

The conditions of integrability of (10.3–15) arise from

$$\boldsymbol{o} = \nabla_\lambda \nabla_\mu \boldsymbol{x} - \nabla_\mu \nabla_\lambda \boldsymbol{x} = \nabla_\lambda \boldsymbol{x}_\mu - \nabla_\mu \boldsymbol{x}_\lambda,$$

when we take account of (10.3–16). We get

$$\boldsymbol{y}(h_{\lambda\mu} - h_{\mu\lambda}) = \boldsymbol{o},$$

an identity, for the functions $h_{\lambda\mu}$ are assumed to be symmetric with respect to their subscripts.

The conditions of integrability of (10.3–16) arise from

$$-\boldsymbol{x}_\kappa R^\kappa{}_{\beta\lambda\mu} = \nabla_\lambda \nabla_\mu \boldsymbol{x}_\beta - \nabla_\mu \nabla_\lambda \boldsymbol{x}_\beta = \boldsymbol{y}(\nabla_\lambda h_{\mu\beta} - \nabla_\mu h_{\lambda\beta}) + \nabla_\lambda \boldsymbol{y}\, h_{\mu\beta} - \nabla_\mu \boldsymbol{y}\, h_{\lambda\beta}.$$

In view of (10.3–17) we have

$$\nabla_\lambda \boldsymbol{y}\, h_{\mu\beta} - \nabla_\mu \boldsymbol{y}\, h_{\lambda\beta} = \boldsymbol{x}_\kappa g^{\alpha\kappa}(-h_{\mu\beta} h_{\lambda\alpha} + h_{\lambda\beta} h_{\mu\alpha}) \\ = -\boldsymbol{x}_\kappa g^{\alpha\kappa} R_{\alpha\beta\lambda\mu} = -\boldsymbol{x}_\kappa R^\kappa{}_{\beta\lambda\mu}.$$

Also in this case we obtain identities, for the equations (10.3–7) are equivalent to $\nabla_\lambda h_{\mu\beta} = \nabla_\mu h_{\lambda\beta}$.

Finally we observe that the conditions of integrability of (10.3–17) arise from

$$\boldsymbol{o} = \nabla_\lambda \nabla_\mu \boldsymbol{y} - \nabla_\mu \nabla_\lambda \boldsymbol{y} = \boldsymbol{x}_\kappa g^{\alpha\kappa}(-\nabla_\lambda h_{\mu\alpha} + \nabla_\mu h_{\lambda\alpha}) + \\ -\nabla_\lambda \boldsymbol{x}_\kappa g^{\alpha\kappa} h_{\mu\alpha} + \nabla_\mu \boldsymbol{x}_\kappa g^{\alpha\kappa} h_{\lambda\alpha} = (\nabla_\mu \boldsymbol{x}_\kappa h_{\lambda\alpha} - \nabla_\lambda \boldsymbol{x}_\kappa h_{\mu\alpha}) g^{\alpha\kappa}.$$

Referring again to (10.3–16) we obtain for this last expression

$$\boldsymbol{y}(h_{\lambda\alpha} h_{\mu\kappa} - h_{\mu\alpha} h_{\lambda\kappa}) g^{\alpha\kappa}$$

and by inspection we see that this expression vanishes identically. Thus we have proved:

Under the assumptions (10.3–7) *and* (10.3–8) *the system consisting of the equations* (10.3–15), (10.3–16) *and* (10.3–17) *is completely integrable.*

As a consequence we can find a uniquely determined set of vector functions $\boldsymbol{x}(q^\kappa), \boldsymbol{y}(q^\kappa), \boldsymbol{x}_\lambda(q^\kappa), \lambda, \kappa = 1, \ldots, n$, taking prescribed values for given values of the independent variables q^κ.

10.3.3 – Solution of Bonnet's problem

Continuing our considerations we may now show that a hypersurface (10.3–3), where $\boldsymbol{x}$ is obtained from the equations (10.3–15), (10.3–16) and (10.3–17), admits the functions (10.3–1) and (10.3–2) as the components of the first and the second fundamental tensor respectively. It is understood, of course, that (10.3–9) also holds throughout a region in the numbers pace of the q^κ. To this end we introduce the functions

$$z_{\lambda\mu} = \boldsymbol{x}_\lambda \boldsymbol{x}_\mu, \qquad z_\lambda = \boldsymbol{x}_\lambda \boldsymbol{y}, \qquad z = \boldsymbol{y}\boldsymbol{y}. \tag{10.3–18}$$

They are solutions of the system

$$\nabla_\mu z_{\beta\gamma} = h_{\beta\mu} z_\gamma + h_{\gamma\mu} z_\beta, \tag{10.3–19}$$

$$\nabla_\mu z_\beta = -h_{\alpha\mu} g^{\alpha\kappa} z_{\kappa\beta} + h_{\beta\mu} z, \tag{10.3–20}$$

$$\nabla_\mu z = -2h_{\alpha\mu} g^{\alpha\kappa} z_\kappa. \tag{10.3–21}$$

The verification that this system is also completely integrable offers no difficulty. Hence there is only one set of solutions with prescribed initial values for given values of the independent variables q^κ. By inspection we see that the functions

$$z_{\beta\gamma} = g_{\beta\gamma}, \qquad z_\gamma = 0 \quad , \quad z = 1, \tag{10.3–22}$$

satisfy the system under consideration. Now it is always possible to find a set of solutions of the system considered in the preceding section, such that its values for given values of q^κ satisfy the conditions (10.3–22). Since the solutions of (10.3–19), (10.3–20) and (10.3–21) are uniquely determined by their initial values we, therefore, have for all allowable values of q^κ:

$$\boldsymbol{x}_\lambda \boldsymbol{x}_\mu = g_{\lambda\mu}, \qquad \boldsymbol{x}_\lambda \boldsymbol{y} = 0, \qquad \boldsymbol{y}\boldsymbol{y} = 1. \tag{10.3–23}$$

On account of (10.3–12) we also have

$$\partial_\lambda \boldsymbol{x} \partial_\mu \boldsymbol{x} = g_{\lambda\mu} \tag{10.3–24}$$

and thus we see that the $g_{\lambda\mu}$ are the components of the first fundamental tensor of the hypersurface $\boldsymbol{x}(q^\kappa)$. If we identify $\boldsymbol{y}$ with the normal $\boldsymbol{n}$ we find from (10.3–13) and (10.3–23) that

$$h_{\lambda\mu} = n\partial_{\lambda\mu}x, \tag{10.3–25}$$

i.e., the functions $h_{\lambda\mu}$ are the components of the second fundamental tensor.

Assume now that we have constructed a second hypersurface $\overset{*}{x}$ such that this and the corresponding functions $\overset{*}{x}_\kappa$, $\overset{*}{y}$ take prescribed values satisfying (10.3–22) for some given values of the variables q^κ. To these values corresponds a point of the manifold and at this point we have

$$\overset{*}{x}_\lambda \overset{*}{x}_\mu = g_{\lambda\mu}, \qquad \overset{*}{x}_\lambda \overset{*}{y} = 0, \qquad \overset{*}{y}\overset{*}{y} = 1. \tag{10.3–26}$$

But when we replace the initial values of $\overset{*}{x}$, $\overset{*}{x}_\kappa$ by the values of x, x_κ for the second values of q^κ, we have moved the second manifold to another which coincides with the first one on account of the uniqueness. The equations (10.3–23) and (10.3–26) express the fact that the frames $(\overset{*}{x}_1, \ldots, \overset{*}{x}_n, \overset{*}{y})$ and $(x_1, \ldots, x_n, y)$ evaluated for the same values of the independent variables are congruent. Thus we arrive at the result:

There exists a hypersurface with prescribed first and second fundamental tensor, provided the equations of Gauss and Codazzi are satisfied. All these hypersurfaces are congruent.

10.3.4 – The equations of applicability

The result obtained in the preceding section enables us to give an answer to the following question: Under what conditions are the functions (10.3–1) the components of the metric tensor of a hypersurface?

This problem may be attacked in the following way. Let $x(q^\kappa)$ denote a hypersurface. At a given point we take an orthonormal set of (tangential) unit vectors $\underset{h}{e}$, $h = 1, \ldots, n$. Combined with the normal n they constitute an orthonormal set in the ambient space. Next we consider a unit vector e in the space, this vector being independent of the parameters q^κ. We may decompose this vector with respect to the orthonormal frame under consideration by writing

$$e = n(e\,n) + \sum_{h=1}^{n} \underset{h}{e}(e\,\underset{h}{e}). \tag{10.3–27}$$

If

$$\underset{h}{e} = x_\kappa \underset{h}{e}^\kappa, \qquad x_\kappa = \partial_\kappa x, \tag{10.3–28}$$

and if we introduce the scalar invariant

$$x = ex \tag{10.3–29}$$

then, evidently,

$$\underset{h}{\boldsymbol{e}}\,\boldsymbol{e} = \underset{h}{e^{\kappa}}\,\partial_{\kappa} x, \qquad h = 1, \ldots, \mathrm{n}. \tag{10.3–30}$$

Hence the norm of the vector (10.3–27) may be expressed as

$$\begin{aligned} 1 &= (\boldsymbol{en})^2 + \sum_{h=1}^{\mathrm{n}} \underset{h}{e^{\lambda}}\,\underset{h}{e^{\mu}}\,\partial_\lambda x \partial_\mu x \\ &= (\boldsymbol{en})^2 + g^{\lambda\mu}\partial_\lambda x \partial_\mu x = (\boldsymbol{en})^2 + \Delta_1(x, x), \end{aligned} \tag{10.3–31}$$

where we have made use of (2.3–18). It follows that

$$\boldsymbol{en} = \pm\sqrt{1-\Delta_1(x, x)}. \tag{10.3–32}$$

Multiplying both members of (10.3–4) by $\boldsymbol{e}$ and introducing the functions

$$x_{\lambda\mu} = \partial_{\lambda\mu} x - \Gamma^{\kappa}{}_{\lambda\mu}\partial_\kappa x \tag{10.3–33}$$

we get

$$x_{\lambda\mu} = h_{\lambda\mu}\boldsymbol{en}. \tag{10.3–34}$$

Finally we obtain from (10.3–8)

$$\boxed{x_{\alpha\lambda}x_{\beta\mu} - x_{\alpha\mu}x_{\beta\lambda} = R_{\alpha\beta\lambda\mu}(1-\Delta_1(x, x)).} \tag{10.3–35}$$

These equations are known as the *equations of applicability.*

Our main result will be the following:

A necessary and sufficient condition for the existence of a hypersurface with prescribed first fundamental tensor is that the system of partial differential equations (10.3–35) *possess a solution with* $\Delta_1(x, x) < 1$.

We only have to verify the sufficiency of the condition stated in the theorem. Let us assume that the system (10.3–35) possesses a solution $x(q^\kappa) \neq 0$. Next we introduce the functions $h_{\lambda\mu}$:

$$h_{\lambda\mu} = \frac{x_{\lambda\mu}}{\sqrt{1-\Delta_1(x, x)}}. \tag{10.3–36}$$

It will be our aim to show that there is a hypersurface with the prescribed metric tensor and having the functions (10.3–36) as components of the second fundamental tensor. To this end we must verify the conditions (10.3–7) and (10.3–8).

In view of (10.3–35) the equations (10.3–8) are valid. But also the verification of (10.3–7) offers no difficulty. We shall write them in the form

$$\nabla_\lambda h_{\beta\mu} = \nabla_\mu h_{\beta\lambda}. \tag{10.3–37}$$

Now it follows from (10.3–36) that

$$\nabla_\lambda h_{\beta\mu} - \nabla_\mu h_{\beta\lambda} \tag{10.3–38}$$
$$= \frac{1}{\sqrt{1-\Delta_1(x, x)}}\left(\nabla_\lambda x_{\beta\mu} - \nabla_\mu x_{\beta\lambda} + \frac{1}{2}\frac{x_{\beta\mu}\nabla_\lambda\Delta_1(x, x) - x_{\beta\lambda}\nabla_\mu\Delta_1(x, x)}{1-\Delta_1(x, x)}\right).$$

In order to prove that the expression between brackets vanishes we observe that in the first place

$$\nabla_\lambda x_{\beta\mu} - \nabla_\mu x_{\beta\lambda} = -R^\kappa{}_{\beta\lambda\mu} x_\kappa.$$

Secondly we have

$$\nabla_\lambda \Delta_1(x, x) = \nabla_\lambda(g^{\alpha\kappa}\partial_\alpha x \partial_\kappa x) = 2g^{\alpha\kappa} x_{\alpha\lambda} x_\kappa,$$

whence

$$x_{\beta\mu}\nabla_\lambda\Delta_1(x, x) - x_{\beta\lambda}\nabla_\mu\Delta_1(x, x)$$
$$= 2g^{\alpha\kappa}(x_{\alpha\lambda}x_{\beta\mu} - x_{\beta\lambda}x_{\alpha\mu}) = 2R^\kappa{}_{\beta\lambda\mu}(1-\Delta_1(x, x)).$$

Inserting these results into (10.3–38) we find (10.3–37). This completes the proof of the theorem.

10.3.5 – The indeformability of a hypersurface

In ordinary space a surface is not determined by its first fundamental tensor alone (even when we leave movements out of consideration). In fact, a sufficiently small part may be deformed in many ways without stretching. Hence there are many second fundamental tensors compatible with the first.

One should expect the same situation in higher dimensional space, since in that case there are more degrees of freedom available. It is, therefore, very surprising that in general a hypersurface is not locally deformable. This fact is stated in the following theorem:

The second fundamental tensor of a hypersurface is determined (within sign) by the first fundamental tensor, when its rank exceeds the number two.

We interpret the second fundamental tensor, evaluated at a preassigned point, as a linear operator defined in an n-dimensional vector space, associated with the tangential hyperplane at the given point. With respect to a given system of parameters the mapping may be represented by

$$u^\kappa = h^\kappa_\lambda x^\lambda. \tag{10.3–39}$$

The dimension of the image space is equal to the rank of the operator with components h^κ_λ (see section 1.2.4) and is, therefore, greater than two. All vector planes through the vector $\boldsymbol{x}$ are transformed into vector planes through the corresponding vector $\boldsymbol{u}$ and they do not coincide with a single vector plane, since the rank of the mapping exceeds two. Hence the vector $\boldsymbol{u}$ is known if the images of the vector planes are known.

A plane is given by a simple bivector (section 4.3.2) and it determines the bivector up to a multiplicative constant. To a simple bivector

$$a^{\kappa\lambda} = x^{\kappa} y^{\lambda} - x^{\lambda} y^{\kappa} \tag{10.3–40}$$

corresponds the bivector

$$b^{\mu\nu} = (h^{\mu}_{\kappa} h^{\nu}_{\lambda} - h^{\mu}_{\lambda} h^{\nu}_{\kappa}) a^{\lambda\kappa} \tag{10.3–41}$$

and by a straightforward calculation we may verify that this bivector is also simple. Hence the mapping of the vector planes is uniquely determined by the tensor

$$R_{\alpha\beta\lambda\mu} = h_{\alpha\lambda} h_{\beta\mu} - h_{\alpha\mu} h_{\beta\lambda}, \tag{10.3–42}$$

where the expression on the left only depends on the components of the first fundamental tensor (section 8.4.3). Since the mapping of the vector planes determines uniquely the mapping of the rays, the components $h_{\lambda\mu}$ are uniquely determined by (10.3–42) within a factor ε. Inserting this factor into (10.3–42) we find $\varepsilon^2 = 1$, proving the theorem.

An equivalent statement of the above theorem is the following:

A hypersurface is locally rigid if more than two of its principal curvatures do not vanish.

In fact, by reducing the operator $h_{\lambda\mu}$ on principal axes, we see that this assertion is equivalent to the statement that the rank exceeds two.

BIBLIOGRAPHY

EISENHART, L. P., Riemannian geometry. Princeton Univ. Press, 1949

———, Non-riemannian geometry. Am. Math. Soc. Coll. Publ., 1927

———, An introduction to differential geometry. Princeton Univ. Press, 1940

HALMOS, P. R., Finite dimensional vector spaces. Princeton Univ. Press, 1953

HOFFMAN, K. and R. KUNZE, Linear Algebra. Prentice Hall, Inc. Englewood Cliffs, N. J., 1961

LAGRANGE, R., Calcul différentiel absolu. Mém. d. Sc. Math. fasc. 19, Gauthier-Villars, Paris, 1926

LEVI CIVITÀ, T., The absolute differential calculus. Blackie and Son, London and Glasgow, 1927

LICHNÉROWICZ, A., Eléments de calcul tensoriel. Colin, Paris, 1950

SCHOUTEN, J. A., Ricci-calculus. An introduction to tensor analysis and its geometrical applications. Springer, Berlin 1954

——— and D. J. STRUIK, Einführung in die neueren Methoden der Differentialgeometrie. Vol. I by J. A. Schouten, Vol. II by D. J. Struik. Noordhoff, Groningen, 1935, 1938

SYNGE, J. L. and A. SCHILD, Tensor calculus. University of Toronto Press, 1956

THOMAS, J. Y., The differential invariants of generalized spaces. Cambridge Univ. Press, 1934

———, Concepts from tensor analysis and differential geometry. Academic Press, New York and London, 1961

VEBLEN, O., Invariants of quadratic forms. Cambr. Tracts no. 24, 1927

——— and C. E. WEATHERBURN, An introduction to Riemannian geometry and the tensor calculus. Cambridge Univ. Press, 1927

WHITEHEAD, J. H. C., The foundations of differential geometry. Cambr. Tracts, no. 29, 1953

WILLMORE, T. J., An introduction to differential geometry. Clarendon Press, Oxford, 1959

INDEX