Wiley Classics Library

STOKER

Differential Geometry

This page intentionally left blank

DIFFERENTIAL GEOMETRY

J. J. STOKER

Wiley Classics Edition Published in 1989

WILEY

WILEY-INTERSCIENCE
a Division of John Wiley & Sons
New York · London · Sydney · Toronto

A NOTE TO THE READER
This book has been electronically reproduced from digital information stored at John Wiley & Sons, Inc. We are pleased that the use of this new technology will enable us to keep works of enduring scholarly value in print as long as there is reasonable demand for them. The content of this book is identical to previous printings.

10 9 8 7 6

Library of Congress Catalog Card Number: 69–16131
SBN 471 82825 4

ISBN 0-471-50403-3 (pbk.)

To Heinz Hopf

This page intentionally left blank

PREFACE

More than thirty-five years ago I was introduced to the subject of this book by my friend and teacher Heinz Hopf through his lectures at the Technische Hochschule in Zürich. I had expected to take my degree there in applied mathematics and mechanics, but Heinz Hopf made such an impression on me, and created such an interest for the subject in me, that I wrote my thesis in differential geometry in the large on a topic suggested by him. My professional career afterwards turned in the main to fields concerned with mathematics in relation to problems in mechanics and mathematical physics generally. However, differential geometry has continued to fascinate me and to cause my thoughts to return again and again to various problems in the large—particularly during the rather frequent occasions when I happened to be teaching a course on the subject. Unfortunately, my efforts in this direction have had rather meagre results, so that I feel myself to be an amateur in the field. However, I am an amateur in the etymological sense of that word, and hope that something of my love for differential geometry will be infectious and will carry over to readers of my book.

In the introduction which follows this preface I outline the contents of the book and indicate the ways in which it differs from others in its attitudes and in its selection of material. In brief, it is stated that the book is intended for students and readers with a minimum of mathematical training, but still has the intention to deal with much that is relatively new in the field, particularly in differential geometry in the large. It also has as one of its purposes the introduction and use of three different notations: vector algebra and calculus; tensor calculus; and the notation devised by Cartan, which employs invariant differential forms as elements in an algebra due to Grassman, combined with an operation called exterior differentiation.

It is now my pleasant duty to thank a number of my friends and colleagues for the help and advice they have given me. Louis Nirenberg and Eugene Isaacson used the manuscript in courses, read it in detail, and spent much time and effort in making specific corrections as well as suggestions of a general character. K. O. Friedrichs also made some use of the manuscript in a course, and I benefited from a number of discussions with him about a variety of matters of principle and logic which aroused his interest. H. Kar-

cher gave me a number of valuable suggestions about some parts of Chapter VIII, which deals with problems in the large; his help is acknowledged in that chapter at the appropriate places.

I owe much to Miss Helen Samoraj, who typed the manuscript in several versions, uncovered many errors and mistakes, and prodded me from time to time to get on with the job. I wish also to thank Carl Bass for drawing the figures.

In 1964 the Guggenheim Foundation gave me a Fellowship; during that time this book was finally organized and carried far toward completion in some of its major portions.

Finally, I am very happy to acknowledge the help given to me by the Mathematics Branch of the Office of Naval Research. I do this with particular pleasure because I have felt for years that the Office of Naval Research has had a very remarkable and beneficial effect on the progress of science in this country.

James J. Stoker
Professor of Mathematics
Courant Institute of
Mathematical Sciences
New York University

INTRODUCTION

Differential geometry is a subject of basic importance for all mathematicians, regardless of their special interests, and it also furnishes essential ideas and tools needed by physicists and engineers. But important as these considerations are, the value of the subject, for the author at least, arises rather from the great variety and beauty of the material itself, and for the close ties it has with important portions of algebra, topology, non-euclidean geometry, analysis generally (in particular with the theory of partial differential equations), and in mechanics and the general theory of relativity. Beside all that it furnishes a great variety of fascinating unsolved problems of its own that are of a particularly challenging nature.

In writing this book the author had in mind these different points of view, and the corresponding classes of potential readers with their various interests. The intention, therefore, is not to present a treatise for advanced students and specialists, but rather to present an introductory book which assumes no more at the outset than a knowledge of linear algebra and of the basic elements of analysis—in other words such preparation as an advanced undergraduate student of mathematics could be expected to have, and the kind of preparation to be expected in the early years of graduate study for the other classes of readers indicated above. It turns out, happily, that even quite recent and interesting advances in the subject can be dealt with on the basis of such relatively scanty foreknowledge.

Since there are quite a number of books about differential geometry in print the author feels it his duty to say in what ways his book differs from others in its attitudes and its selection of material. A brief outline of the contents of the book, chapter by chapter, is therefore given here.

Chapter I gives a brief summary of the basic facts and notations of vector algebra and calculus that are used in the book.

Chapter II deals with the theory of regular curves in the plane. Most books, if they deal with plane curves at all, consider them as a special case of space curves. In this book a relatively long chapter is devoted to them because they are of great interest in their own right and their theory is not in all respects the same as it is for space curves. In addition, it is possible to present the theory of plane curves in such a way as to give the basic general motivations once for all for the underlying concepts of differential geometry so that the concepts can be introduced without much motivation in the more complicated cases of space curves and surfaces. Included in Chapter II is a discussion of some problems in differential geometry in the large. Among them is a proof of the Jordan theorem for smooth plane curves having a uniquely determined tangent vector. This belongs in differential geometry, whereas the theorem for merely continuous curves properly belongs in topology. The author found no proof in the literature of the Jordan theorem including a proof of the fact that the interior domain is simply connected, for the simpler case of smooth curves, in spite of the fact that the theorem in this form is the most widely used—for example in the integration theory for analytic functions of a complex variable, and in mechanics.

Chapter III deals with the theory of twisted curves in three space. The concepts of arc length s, curvature κ, and torsion τ are introduced. The Frenet equations are derived, and on the basis of the existence and uniqueness theorems for ordinary differential equations it is shown that the three invariants s, κ, and τ form a complete set of invariants in the sense that any two curves for which these quantities are the same differ at most by a rigid motion. Important connections of this theory with the kinematics of rigid body motion, and of the motion of a particle under given forces, are discussed.

Chapter IV deals with the basic elements of the theory of regular surfaces in three dimensional space. This revolves to a large extent around the two fundamental quadratic differential forms which serve to define the length of curves on the surface and the various curvatures that can be defined on it. Interesting special curves such as the asymptotic lines and lines of curvature, and their properties, are studied. The solution of many problems in differential geometry (and in other disciplines as well) can often be made very simple once an appropriate special system of curvilinear coordinates is introduced. The author thought it reasonable to justify such procedures in a number of important cases by an appeal to the existence and uniqueness theorems for ordinary differential equations.

Chapter V is concerned with two special classes of surfaces that are very interesting in their own right and that also serve to illustrate how the theory of Chapter IV can be used. These are the surfaces of revolution and the developable surfaces. The old-fashioned classification of the developables as cylinders, cones, or tangent surfaces of space curves is given up since this classification rules out many valid developables, i.e. many easily defined surfaces that are not composed entirely of parabolic points. Instead these surfaces are defined as those for which the Gaussian curvature is everywhere zero, and various properties of them in the large are treated on this basis.

Chapter VI treats the fundamental partial differential equations of the theory of surfaces in three-space. These come about by expressing the *first* *derivatives* of the two tangent vectors of the coordinate curves on the surface, and of its unit normal vector, as linear combinations of these vectors themselves. These equations are given the names of Gauss and Weingarten. They form an over-determined system, i.e. there are many more equations than there are dependent functions to be determined. Solutions thus exist only when certain compatibility conditions are satisfied, and these conditions are equations due to Gauss and to Codazzi and Mainardi. The equation due to Gauss embodies what is perhaps the most striking theorem in the whole subject: it says that the Gaussian curvature, defined originally for a surface in three-space, is really independent of the form of the surface in three-space so long as the lengths of all curves on the surface remain unchanged in any deformation of it. This theorem gave rise in Gauss's mind to the fruitful idea—later on developed in full generality by Riemann—of dealing with inner differential geometry, i.e. to geometrical questions that concern only geometry in the surface as evidenced by the nature of the length measurements on it. In this kind of geometry all geometric notions arise from the functions which, as its coefficients, serve to define the first fundamental form; much of the later portions of the book are concerned with such inner, or intrinsic, geometries. In Chapter VI it is shown that a surface exists and is uniquely determined within rigid motions once the coefficients of the two fundamental forms are given, and if these functions satisfy the compatability conditions; this is done by integration of the basic partial differential equations. The theorem follows from a basic theorem concerning over-determined systems (a theorem proved in Appendix B).

Chapter VII has as its purpose a treatment of *inner* differential geometry of surfaces, but it is done nevertheless by considering the surfaces to lie in three-space. In this way an intuitive geometric motivation for the concepts of the inner, or intrinsic, geometry of surfaces is made direct and simple. The concept chosen as basic for the whole chapter is the beautiful one due to Levi-Civita, of the parallel transport of a vector along a given curve on the surface. From this the notion of the geodetic curvature, denoted by κ_g, of a given curve is derived, and the special curves called geodetic lines are defined as those for which κ_g vanishes. All of these concepts, though derived for surfaces in three-space, are seen to belong to the intrinsic geometry of surfaces since they make use in the end only of quantities that are completely determined by the coefficients of the first fundamental form. Nevertheless it is quite interesting to know their relation to surfaces embedded in three-space. The geodetic lines, though defined initially as those curves along which $\kappa_g = 0$, can also be defined through studying curves of shortest length between pairs of points on a surface. This important problem is treated partially in Chapter VII. It is shown that the condition $\kappa_g = 0$ (which really is a second order ordinary differential equation) is in general only a *necessary* condition in order that a geodetic line should be a curve of shortest length. On the other hand it is shown that a small enough neighborhood of any point p can always be found such that any point q of it can be joined to p by a uniquely determined geodetic line of shortest length, when compared with the length of any other curve joining p and q. Since the differential equation determining the geodesics is of second order it follows that a uniquely determined geodesic exists through a given point p in every direction. In fact, a certain neighborhood of p is covered simply by these arcs, which can be taken, together with their orthogonal trajectories, as a regular parameter system—in complete analogy with polar coordinates in the plane or spherical coordinates on the sphere. This is one of those special coordinate systems referred to earlier that have the effect of simplifying the solutions of particular problems. One such problem concerns the surfaces of constant Gaussian curvature K, which are seen to furnish models for the three classical geometries, i.e. the Euclidean for $K = 0$, the elliptic geometry for $K > 0$, and the hyperbolic or Lobachefsky geometry for $K < 0$ when the straight lines are defined as geodesics in their whole extent. The Lobachefsky geometry is treated in some detail. The Gauss-Bonnet formula is derived (in the small) in this chapter. This formula relates the integral of the Gaussian curvature over a simply connected domain to the integral of the geodetic curvature over the boundary curve of the domain. A tool used to accomplish this is also derived; it is the beautiful result that the integral of the Gaussian curvature over a domain is equal to the angle change that results when a vector is transported parallel to itself around the boundary of the domain.

Chapter VIII is probably the chapter that makes the book most different from others because it deals with a considerable variety of the fascinating theorems of differential geometry in the large, especially for two-dimensional manifolds. For this purpose an introduction to the concept of a manifold in n dimensions is given intrinsically. This leads to the special case of a Riemannian manifold. Since most of the material of the chapter is then specialized to two-dimensional manifolds—in fact in large part to the compact two-dimensional manifolds—it was thought reasonable to interpolate a brief description of the facts from topology about them that are needed later on. Except for the last two sections of the chapter the theorems in the large are all concerned with inner differential geometry, thus indicating that this kind of geometry is very rich in content. Once abstract surfaces or manifolds have been given a metric it is possible to consider them in a natural way as metric spaces by defining a distance function in them, and to introduce the concept of *completeness*. This means, roughly speaking, that the manifold contains no boundary points at finite distance from any given point; thus this condition is a restriction only for open, or non-compact, manifolds. The theorem of Hopf and Rinow, which establishes the equivalence of four different characterizations of completeness, is proved as a by-product of the proof of one of the most important single theorems in differential geometry in the large, i.e. the theorem that a curve of shortest length exists joining any pair of points on a complete manifold, and that this curve is a geodetic line. A section is devoted to angle comparison theorems, of rather recent date, for geodetic triangles on surfaces. Geodetically convex domains are studied; in particular it is shown that sufficiently small geodetic circles and geodetic triangles are geodetically convex. The Gauss-Bonnet formula is used to prove the beautiful theorem that the integral of the Gaussian curvature over the area of a two-dimensional compact surface is not only an isometric invariant, but is also a topological invariant with a value fixed by the Euler characteristic. Vector fields on surfaces are considered and an index is assigned to their isolated singularities, i.e. to points where the field vector is the zero vector. This makes it possible to prove a theorem due to Poincaré, with the aid of the theorem on the change of angle resulting by parallel transport of a vector around a simple closed curve, that determines the sum of the indices in question on a compact surface in terms of its Euler characteristic. The theorem on the existence of shortest arcs as geodesics referred to earlier was a nonconstructive existence theorem. It is of great interest to approach this problem more directly as a two-point boundary value problem for the second order ordinary differential equation that characterizes geodetic lines. This leads to Jacobi's theory of the second variation and to a sufficient condition, based on the notion of a conjugate point, for the existence of the shortest join when the comparison curves are restricted to a neighborhood of a geodesic joining two points. This theory in turn makes it possible to prove the generalization of a famous theorem of Bonnet given by Hopf and Rinow, i.e. that a complete two-dimensional surface with Gaussian curvature above a certain positive bound is of necessity compact, because it has a diameter that can be estimated in terms of the bound on the curvature, and consequently is readily seen to be topologically a sphere. (Bonnet *assumed* the surface to be topologically a sphere lying in three-space, and then gave a bound for its diameter in terms of the bound on the Gaussian curvature.) The theorem of Synge is next dealt with; this theorem states that a compact manifold with an even number of dimensions and with positive Gaussian curvature is simply connected. The consideration of problems in intrinsic geometry in the large ends with a discussion of covering surfaces of complete two-dimensional surfaces with nonpositive Gaussian curvature—they are obtained by expanding geodetic polar coordinates over the surface. The final two sections of the chapter are concerned with complete surfaces lying in three-dimensional space. The first of these deals with Hilbert's famous theorem on the nonexistence in three-space of a complete regular surface with constant negative Gaussian curvature. Two proofs of the theorem are given. One of them is a version of Hilbert's original proof, but it makes use of the covering surface just mentioned above; the other is a version of a proof due to Holmgren. The final section treats a generalization of a theorem due to Hadamard. This theorem states that a compact surface in three-space with positive Gaussian curvature is the full boundary of a convex body, or, in other words it is an ovaloid. Thus local convexity of the surface combined with its assumed closure guarantees that no double points or self-intersections can occur—in contrast with what can occur for locally convex closed curves in the plane. The theorem is generalized to the case of complete surfaces in three-space, with the result that the open, that is, noncompact, surfaces are the full boundary of an unbounded convex body.

Chapter IX treats the elements of Riemannian geometry on the basis of a systematic, though brief, introduction to tensor calculus. The point of view of Cartan is taken in doing this. Some applications to problems in the large are made, e.g. the extensions of the theorem of Hopf and Rinow, and of Synge's theorem, to n-dimensional manifolds are treated. Although it might be thought to fall out of the scope of a book on differential geometry to treat the general theory of relativity, the author nevertheless thought it good to do that. The reason is simply that this application of Riemannian geometry is so striking and beautiful, and it lends itself to a not too lengthy treatment, on a somewhat intuitive basis, even when the special theory of relativity and the relativistic dynamics of particle motion are first explained.

Chapter X has two purposes in view. One of them is to introduce still another notation to those already used. It is a notation due to Cartan which applies an algebra introduced by Grassman, and which employs an alternating product, to elements that are invariant differential forms. (They are invariant by virtue of the fact that their coefficients are the components of alternating covariant tensors.) In addition, an operation called exterior differentiation is introduced. This leads to the construction of new invariant differential forms, of higher degree, from any given one. It turns out that this notation is particularly effective in dealing with compatibility conditions and in converting volume integrals into surface integrals with the use of Green's theorem—both of which are basic operations in differential geometry. The notation is then applied here to vector differential forms in order to formulate the geometry of two-dimensional surfaces in three-space. The compactness of the notation is rather remarkable. In particular, the derivation of such a basic theorem as that of the isometric invariance of the Gaussian curvature, is very elegant. Minimal surfaces are treated here. However, most of the applications treated in this chapter are concerned with differential geometry in the large. These include various characterizations of the sphere (Chern's theorem), and three classic theorems concerning the uniqueness within motions of closed convex surfaces in three-space. The three theorems prove the uniqueness of the surface when (1) the line element is prescribed, or (2) the sum of the principal radii of curvature is prescribed as a function of the direction of the surface normal (Christoffel's theorem), or (3) the same as (2) but the Gaussian curvature is prescribed (Minkowski's theorem). These problems are all solved with the aid of an appropriately chosen invariant scalar differential form which results by taking a scalar triple product of three vector differential forms that involve vectors from both of two examples of the surfaces satisfying the given conditions.

A number of problems are formulated at the end of the chapters. The author tried to invent some new problems to serve as exercises; it is hoped that they will be found interesting and instructive without being too difficult.

The book has two appendices. Appendix A summarizes the main facts and formulas needed from linear algebra in a form suitable for ready reference in the book, together with brief discussions of geometry in affine, Euclidean, and Minkowskian spaces. Appendix B gives brief formulations, without proofs, of the basic existence and uniqueness theorems for ordinary differential equations, and a proof of the existence and uniqueness theorems—of such vital importance in differential geometry—for the solutions of over-determined systems of partial differential equations when appropriate compatibility conditions are satisfied.

This outline of the contents of the book should support the earlier statement concerning the author's intentions, i.e. to write (1) a thorough but elementary treatment of differential geometry for young students, that (2) includes a treatment of a rather large number of problems of differential geometry in the large, and that (3) makes a point of introducing and using three different notations employing vectors, then tensors, and finally invariant differential forms. In addition, it is hoped that all of these things can be done successfully on the basis of a minimum of preparation in other mathematical disciplines.

This page intentionally left blank

CONTENTS

Chapter III Space Curves

Chapter IV The Basic Elements of Surface Theory

This page intentionally left blank

CHAPTER I

Operations with Vectors

1 The Vector Notation

This chapter presents briefly the principal rules for operating with vectors, and a collection of those formulas which are useful in differential geometry. No attempt at completeness nor at an axiomatic treatment of vector algebra is made—for that, the student should consult the books about linear algebra (e.g., the book of Gelfand [G.2]), however, a summary of those parts of linear algebra that are most relevant to differential geometry is included as Appendix A of this book. In any case, only vector algebra and the elements of vector calculus are needed in the first eight chapters. Later on in Chapter IX the tensor calculus, and in Chapter X the notation based on invariant forms and their exterior derivatives, will be introduced and applied.

Vectors are denoted by Latin letters in bold-faced type, usually as capital letters, except for the case of unit coordinate vectors, which will be denoted by small letters. The rectangular components of a given vector, which are, of course, scalars, will be represented by the corresponding small letter with a subscript:

$$\mathbf{X} = (x_1, x_2, x_3). \tag{1.1}$$

It is often convenient to work with the representation in terms of components; in general, as (1.1) indicates, the coordinate axes will be denoted by x_1, x_2, x_3, as in Fig. 1.1, and they will be chosen so as to form a right-handed coordinate system. The components of the vector are also the coordinates of its end point, the initial point being the origin. By the length $|\mathbf{X}|$, or magnitude, of a vector we mean the length of the straight-line segment from the origin to the point with the coordinates x_1, x_2, x_3; thus we have

$$|\mathbf{X}| = \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{1.2}$$

as the definition for the magnitude of $\mathbf{X}$.

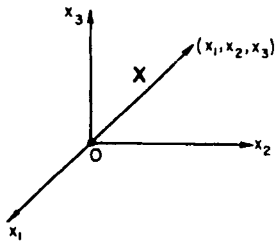

Fig. 1.1 A vector and its components.

2. Addition of Vectors

The characteristic property of vectors that distinguishes them from scalars is embodied in the law of addition, which is the familiar parallelogram law, as indicated in Fig. 1.2. We write

$$\mathbf{Z} = \mathbf{X} + \mathbf{Y} = \mathbf{Y} + \mathbf{X}. \tag{1.3}$$

The order in which the vectors are added is immaterial. Also, the ordinary plus sign is used to denote vector addition. It should be stated explicitly that *vectors can be added in general only when they are attached to the same point.* (In the kinematics and mechanics of rigid bodies certain special types of vectors are not thus restricted, but that is a very exceptional state of affairs.)

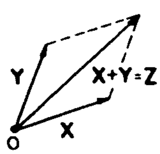

Fig. 1.2 Addition of vectors.

In a sum of several vectors parentheses may be introduced or taken away at will:

$$\mathbf{X} + (\mathbf{Y} + \mathbf{Z}) = (\mathbf{X} + \mathbf{Y}) + \mathbf{Z} = \mathbf{X} + \mathbf{Y} + \mathbf{Z}. \tag{1.4}$$

In terms of the representation using components, the rule (1.3) reads

$$\mathbf{Z} = (x_1 + y_1, x_2 + y_2, x_3 + y_3) = (y_1 + x_1, y_2 + x_2, y_3 + x_3). \tag{1.5}$$

3. Multiplication by Scalars

Various different sorts of products occur in vector algebra. Consider first the product $\alpha\mathbf{X}$ of a scalar and a vector; this means geometrically that the length, or magnitude, of $\mathbf{X}$ is multiplied by α, but the direction is either left unaltered (if $\alpha > 0$) or reversed (if $\alpha < 0$). If $\alpha = 0$, the result is the vector zero, which is, however, not printed in bold-faced type since no confusion will result in this exceptional case. In terms of the components of $\mathbf{X}$ the product $\alpha\mathbf{X}$ is given by

$$\alpha\mathbf{X} = (\alpha x_1, \alpha x_2, \alpha x_3). \tag{1.6}$$

In this notation the fact that $\alpha\mathbf{X}$ is opposite in direction to $\mathbf{X}$ for α negative is clear. It is also clear that a difference of two vectors is to be interpreted as the sum of $\mathbf{X}$ and of -1 times $\mathbf{Y}$, or, as it is also put as the sum of $\mathbf{X}$ and of the vector obtained by reversing the direction of $\mathbf{Y}$.

The following rules hold for the product of a scalar and a vector:

$$\begin{gathered}(\alpha + \beta)\mathbf{X} = \alpha\mathbf{X} + \beta\mathbf{X}, \qquad \alpha(\mathbf{X} + \mathbf{Y}) = \alpha\mathbf{X} + \alpha\mathbf{Y} \\ \alpha(\beta\mathbf{X}) = (\alpha\beta)\mathbf{X} = \alpha\beta\mathbf{X}\end{gathered} \tag{1.7}$$

4. Representation of a Vector by Means of Linearly Independent Vectors

An important fact about vectors in the three-dimensional Euclidean space is that any vector $\mathbf{V}$ can be expressed in one, and only one, way as a linear combination of any three vectors $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ which do not lie in the same plane; that is, uniquely determined scalars α, β, γ exist under these circumstances such that

$$\mathbf{V} = \alpha\mathbf{X} + \beta\mathbf{Y} + \gamma\mathbf{Z}. \tag{1.8}$$

Three vectors $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ that are not at all in the same plane are said to be linearly independent.

In two dimensions, that is, in the plane, any vector can be expressed as a linear combination of any two others which are not in the same straight line; again it is said that the vector is expressed as a linear combination of linearly independent vectors.

5. Scalar Product

Another kind of product, called a scalar product, involves the multiplication of two vectors, but in such a manner as to yield a scalar quantity. The notation for this product is $\mathbf{X}\cdot\mathbf{Y}$; it is in fact sometimes called the dot product of the vectors. It is defined as follows:

$$\mathbf{X}\cdot\mathbf{Y} = |\mathbf{X}|\,|\mathbf{Y}|\cos\theta, \tag{1.9}$$

in which θ is the angle, $0 \le \theta \le \pi$, between the two vectors, as shown in Fig. 1.3. It is the product of the lengths of the two vectors and the cosine of the

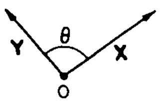

Fig. 1.3 The scalar product.

angle between them. It is also the product of the length of either one of the vectors and the length of the projection of the other vector on it. The following rules for operating with this product hold:

$$\begin{aligned} \mathbf{X}\cdot\mathbf{Y} &= \mathbf{Y}\cdot\mathbf{X}, \\ \mathbf{X}\cdot(\mathbf{Y}+\mathbf{Z}) &= \mathbf{X}\cdot\mathbf{Y} + \mathbf{X}\cdot\mathbf{Z}, \\ (\alpha\mathbf{X})\cdot\mathbf{Y} &= \alpha(\mathbf{X}\cdot\mathbf{Y}) = \alpha\mathbf{X}\cdot\mathbf{Y}. \end{aligned} \tag{1.10}$$

The special case in which

$$\mathbf{X}\cdot\mathbf{Y} = 0 \tag{1.11}$$

is quite important; this equation holds not only if $\mathbf{X}$ or $\mathbf{Y}$ is zero but also if neither $\mathbf{X}$ nor $\mathbf{Y}$ is the zero vector but the two are *orthogonal*. We observe also that

$$\mathbf{X}\cdot\mathbf{X} = |\mathbf{X}|^2. \tag{1.12}$$

The scalar product of a vector with itself thus gives the square of the magnitude of the vector. Sometimes $\mathbf{X}\cdot\mathbf{X} = \mathbf{X}^2$ is written if there is no danger of ambiguity.

Consider an orthogonal right-handed coordinate system with vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ along the coordinate axes, with $|\mathbf{u}_i| = 1$ (i.e., these are so-called *unit* vectors). Any vector can be represented in the form [cf. (1.8)]

$$\mathbf{X} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + x_3\mathbf{u}_3.$$

In this case the scalars x_i are at once seen to be the components of $\mathbf{X}$. Take also another vector $\mathbf{Y}$ expressed in the same form:

$$\mathbf{Y} = y_1\mathbf{u}_1 + y_2\mathbf{u}_2 + y_3\mathbf{u}_3.$$

The scalar product of the two vectors can be expressed in terms of the components x_i and y_i simply by using the rules given in (1.10); the result is

$$\mathbf{X}\cdot\mathbf{Y} = x_1y_1 + x_2y_2 + x_3y_3, \tag{1.13}$$

since

$$\mathbf{u}_i\cdot\mathbf{u}_j = \delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases} \tag{1.14}$$

these last being relations which hold for any system of mutually orthogonal unit vectors. The convenient and much used symbol δ_{ij}, called the Kronecker delta, is introduced in (1.14). A special case furnishes the well-known relation for the square of the magnitude of a vector:

$$\mathbf{X}\cdot\mathbf{X} = |\mathbf{X}|^2 = x_1^2 + x_2^2 + x_3^2. \tag{1.15}$$

6. Vector Product

Another type of product involving two vectors will be much used. It is a product which yields a new *vector*, and not a scalar, in contrast with the above defined scalar product. The vector product of **X** and **Y** is a *vector* **Z** defined as follows (cf. Fig. 1.4):

$$\mathbf{X} \times \mathbf{Y} = \mathbf{Z} = (|\mathbf{X}|\,|\mathbf{Y}| \sin \theta)\mathbf{u}, \tag{1.16}$$

in which **u** is a unit vector perpendicular to both **X** and **Y** and so taken that the vectors **X**, **Y**, **u**, in that order, form a right-handed system. It is important to observe that $\mathbf{X} \times \mathbf{Y} = -\mathbf{Y} \times \mathbf{X}$, i.e., the vector product is not commutative. Note also that the vector $\mathbf{X} \times (\mathbf{Y} \times \mathbf{Z})$ is not in general the

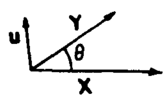

Fig. 1.4 The vector product.

same as the vector $(\mathbf{X} \times \mathbf{Y}) \times \mathbf{Z}$, since the first is in the plane of **Y** and **Z**, the second in the plane of **X** and **Y**. The following rules involving this product can be established with no great difficulty:

$$\begin{aligned} \mathbf{X} \times (\mathbf{Y} + \mathbf{Z}) &= \mathbf{X} \times \mathbf{Y} + \mathbf{X} \times \mathbf{Z}, \\ (\alpha\mathbf{X} \times \mathbf{Y}) &= \alpha(\mathbf{X} \times \mathbf{Y}) = \alpha\mathbf{X} \times \mathbf{Y}. \end{aligned} \tag{1.17}$$

Note that $\mathbf{X} \times \mathbf{Y}$ furnishes the area, with a certain orientation, of the parallelogram determined by **X** and **Y**. In speaking, this product is read "**X** cross **Y**," and, indeed, it is often referred to as the cross product.

As with the scalar product, the special case

$$\mathbf{X} \times \mathbf{Y} = 0, \tag{1.18}$$

in which the vector product vanishes, is important. It occurs, clearly, if either **X** or **Y** is the zero vector, but also if **X** and **Y** fall in the same straight line, that is, if **X** and **Y** are linearly dependent. In particular, it is always true that

$$\mathbf{X} \times \mathbf{X} = 0, \tag{1.19}$$

a formula which comes into play rather often.

The vector product of two vectors **X** and **Y**, when each is represented as a linear combination of a set $\mathbf{u}_i$ of orthogonal unit vectors forming a right-handed system, is readily calculated. The rules in (1.17) can be used to obtain this product in the form

$$\mathbf{X} \times \mathbf{Y} = \mathbf{u}_1(x_2y_3 - x_3y_2) + \mathbf{u}_2(x_3y_1 - x_1y_3) + \mathbf{u}_3(x_1y_2 - x_2y_1), \tag{1.20}$$

when it is observed that $\mathbf{u}_i \times \mathbf{u}_i = 0$ and that $\mathbf{u}_i \times \mathbf{u}_j = \pm\mathbf{u}_k$, the sign depending upon whether or not j follows i in the order $1 - 2 - 3 - 1$. A useful way to remember the formula (1.20) is to put it in the form

$$\mathbf{X} \times \mathbf{Y} = \begin{vmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}, \tag{1.21}$$

which, if developed as though it were an ordinary determinant, leads to (1.20).

The vector product, unlike the scalar product, is not invariant under all orthogonal transformations of the coordinates, but rather is seen to change sign if the orientation of the coordinate axes is changed.

7. Scalar Triple Product

Finally, it is useful to introduce and discuss a special type of product involving three vectors that is defined by the formula $(\mathbf{X} \times \mathbf{Y})\cdot\mathbf{Z}$. That is, the scalar product of **Z** is taken with the vector product of **X** and **Y**; it is called the mixed product, or scalar triple product. As can be read from Fig. 1.5, it represents the volume (with a definite sign) of the parallelepiped, with the three vectors determining its edges. The sign of the product is positive if **X**, **Y**, **Z**, in that order, form a right-handed system of vectors; otherwise the sign is negative.

The following formulas hold:

$$(\mathbf{X} \times \mathbf{Y})\cdot\mathbf{Z} = (\mathbf{Y} \times \mathbf{Z})\cdot\mathbf{X} = (\mathbf{Z} \times \mathbf{X})\cdot\mathbf{Y}, \tag{1.22}$$

but

$$(\mathbf{Y} \times \mathbf{X})\cdot\mathbf{Z} = -(\mathbf{X} \times \mathbf{Y})\cdot\mathbf{Z}.$$

From the second expression in the first line and the fact that $\mathbf{X}\cdot\mathbf{V} = \mathbf{V}\cdot\mathbf{X}$ it is clear that $\mathbf{X}\cdot(\mathbf{Y} \times \mathbf{Z}) = (\mathbf{X} \times \mathbf{Y})\cdot\mathbf{Z}$, so that dot and cross may be interchanged. In fact, there can be no ambiguity in omitting the parentheses altogether, since the vector product is defined only for two vectors. Thus $\mathbf{X}\cdot\mathbf{Y} \times \mathbf{Z}$ must mean $\mathbf{X}\cdot(\mathbf{Y} \times \mathbf{Z})$.

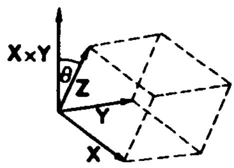

Fig. 1.5 The scalar triple product as a volume.

A useful fact can now be stated: three vectors are linearly independent (and thus span the space) if, and only if, their scalar triple product does not vanish. Or, phrased differently, a necessary and sufficient condition that three vectors should lie in a plane, and thus be linearly dependent, is that the scalar triple product of them should vanish. These and other statements about the scalar triple product can be verified by expressing the three vectors in terms of a system of orthogonal unit coordinate vectors. It is found easily that the triple product is given by the following determinant, the elements of which are the components of the vectors in a right-handed coordinate system:

(1.23) $$\mathbf{X}\cdot\mathbf{Y} \times \mathbf{Z} = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}.$$

8. Invariance Under Orthogonal Transformations

A large part of this book is concerned with the geometry of curves and surfaces which are located in the Euclidean plane or in Euclidean three-space. It is clear that a property of a curve or surface which is entitled to be called a geometrical property must be independent of the special choice of a coordinate system in the space; or, expressed in a different way, such a property should be an invariant under orthogonal linear transformations of the coordinates. The vector notation is well suited for the detection of such properties, since a vector by definition is such an invariant. The scalar product defined above is also an invariant under orthogonal transformations, as one could easily check by a calculation, but which is also obvious from its geometrical interpretation. The vector product is an invariant only under those orthogonal transformations which preserve the orientation of the axes; it changes sign if the orientation is changed. These facts are again easily verified by a calculation, and they are also obvious from the geometrical interpretation of the vector product. The scalar triple product also is an invariant only if the orientation of the coordinate axis is preserved. In general the geometrical properties of curves and surfaces will be defined in terms of vectors, together with the various products of them; thus the invariant character of these properties will be evident.

It might be added that the course pursued in this book eventually leads, in a quite natural way, through the study of the inner geometry of surfaces, to the consideration of geometrical properties that are invariant with respect to more general transformations. At that time the introduction of a more general notation than the vector notation—the tensor notation, for example—becomes a necessity.

When dealing with curves and surfaces in Euclidean space it is natural to speak of *invariance with respect to rigid* motions, and this will sometimes be done. This notion of invariance is conceptually different from that of invariance with respect to transformations of coordinates in the space. By a rigid motion is meant a change of position of an object in the space that preserves the distance between each pair of its points. However, as is well known (see, for example, Appendix A for a discussion of various matters of this kind), such a motion can be described in Euclidean geometry by a mapping of the whole space on itself that preserves distances, and this in turn is achieved by an appropriate orthogonal transformation. Thus, in the end, the two conceptually different notions of invariance both refer to invariance with respect to orthogonal transformations: in the one case with respect to a linear transformation of the whole space into itself, in the other to a transformation of the coordinate system of the space regarded as fixed.

9. Vector Calculus

The vectors dealt with in differential geometry will depend in general upon one or more real scalar parameters. This means simply that the components x_i of the vectors are functions of the parameters. For example, the end point of the vector

$$\mathbf{X}(t) = (x_1(t), x_2(t), x_3(t))$$

will in general fill out a segment of a curve in three-dimensional space when the parameter t varies, as indicated in Fig. 1.6; evidently this is nothing but a short-hand notation which gives the equations of the curve segment in the parametric form $x_i = x_i(t)$, $i = 1, 2, 3$. The vector function $\mathbf{X}(t)$ is said to be *continuous* in $\alpha \leq t \leq \beta$ if the functions $x_i(t)$ are defined and continuous over the interval. The vector $\mathbf{X}(t)$ is said to be *differentiable* if that is true of the coordinates $x_i(t)$, and the derivative of it is defined by the expression

$$\frac{d\mathbf{X}(t)}{dt} = \mathbf{X}'(t) = \left(\frac{dx_1}{dt}, \frac{dx_2}{dt}, \frac{dx_3}{dt}\right), \tag{1.24}$$

or, in terms of a $\mathbf{u}_i$-system of orthogonal unit coordinate vectors, by

$$\mathbf{X}'(t) = x_1'\mathbf{u}_1 + x_2'\mathbf{u}_2 + x_3'\mathbf{u}_3. \tag{1.25}$$

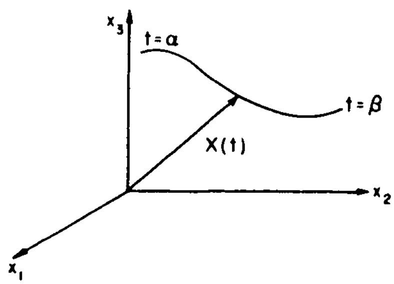

Fig. 1.6 Space curve given by a vector $\mathbf{X}(t)$.

Later on it will become clear why this definition for the derivative of a vector with respect to a scalar parameter is reasonable and appropriate. Here it perhaps suffices to notice that the vector $\mathbf{X}'(t)$ is in the direction which it is customary to define as the direction of the tangent to the curve represented by $\mathbf{X}(t)$.

It is easy to verify that the following rules for differentiation of the various products hold:

$$\begin{aligned}
&\frac{d}{dt}(\alpha(t)\,\mathbf{X}(t)) = \alpha'\mathbf{X} + \alpha\mathbf{X}',\\
&\frac{d}{dt}(\mathbf{X}\cdot\mathbf{Y}) = \mathbf{X}'\cdot\mathbf{Y} + \mathbf{X}\cdot\mathbf{Y}',\\
&\frac{d}{dt}(\mathbf{X}\times\mathbf{Y}) = \mathbf{X}'\times\mathbf{Y} + \mathbf{X}\times\mathbf{Y}',\\
&\frac{d}{dt}(\mathbf{X}\cdot\mathbf{Y})\times\mathbf{Z} = \mathbf{X}'\cdot\mathbf{Y}\times\mathbf{Z} + \mathbf{X}\cdot\mathbf{Y}'\times\mathbf{Z} + \mathbf{X}\cdot\mathbf{Y}\times\mathbf{Z}'.
\end{aligned} \tag{1.26}$$

One caution should be given: the order of the factors must be strictly observed whenever the vector product is involved.

Integration of a vector with respect to a parameter over the range $\alpha \leq t \leq \beta$ is defined, as might be expected, as follows:

$$\int_\alpha^\beta \mathbf{X}(t)\,dt = \mathbf{u}_1 \int_\alpha^\beta x_1(t)\,dt + \mathbf{u}_2 \int_\alpha^\beta x_2(t)\,dt + \mathbf{u}_3 \int_\alpha^\beta x_3(t)\,dt. \tag{1.27}$$

Also, if the upper limit is variable so that the process of integration yields a vector $\mathbf{Y}(t)$ given by

$$\mathbf{Y}(t) = \int_\alpha^t \mathbf{X}(\tau)\,d\tau, \tag{1.28}$$

it follows immediately from (1.27) that

$$\frac{d\mathbf{Y}(t)}{dt} = \mathbf{Y}'(t) = \mathbf{X}(t). \tag{1.29}$$

In other words the analog of the so-called fundamental theorem of the calculus holds for vectors once the above definitions are given for the derivative and the definite integral.

It is frequently useful to employ an analog of the mean value theorem for differentiable scalar functions to a vector function $\mathbf{X}(t)$. Consider, for example, the difference $\mathbf{X}(t_1) - \mathbf{X}(t_0)$, which is given, from (1.5), in terms of the components of $\mathbf{X}(t)$ by

$$\begin{aligned}\mathbf{X}(t_1) - \mathbf{X}(t_0) &= (x_1(t_1) - x_1(t_0), x_2(t_1) - x_2(t_0), x_3(t_1) - x_3(t_0)) \\ &= (x_1'(\xi_1), x_2'(\xi_2), x_3'(\xi_3))\cdot(t_1 - t_0).\end{aligned}$$

The second line is a consequence of the mean value theorem which says that $x_i(t_1) - x_i(t_0) = x_i'(\xi_i)(t_1 - t_0)$ for some value ξ_i between t_0 and t_1. It is often convenient to write this expression in the form

$$\mathbf{X}(t_1) - \mathbf{X}(t_0) = \overset{*}{\mathbf{X}}{}'\cdot(t_1 - t_0), \tag{1.30}$$

with $\overset{*}{\mathbf{X}}{}'$ a vector having the property

$$\lim_{t_1 \to t_0} \overset{*}{\mathbf{X}}{}' = \mathbf{X}'(t_0), \tag{1.31}$$

which holds since the three quantities ξ_i all lie between t_0 and t_1 and the derivatives $x_i'(t)$ are assumed to be continuous. Thus while there is no mean value theorem for vector functions in the same sense as there is for scalar functions of a real variable, the application of that theorem to the components separately leads formally to the relations (1.30) and (1.31) and they can be used in analysis, as will be seen, in much the same fashion as the corresponding formulas are used for scalar functions. The earmark of this procedure in what follows in this book is the star over a derivative of a vector function.

PROBLEMS

1. In the equations $\mathbf{X}\cdot\mathbf{Y} = \alpha$, $\mathbf{X} \times \mathbf{Y} = \mathbf{Z}$ is $\mathbf{Y}$ uniquely determined if $\mathbf{X}$ and α or $\mathbf{X}$ and $\mathbf{Z}$ are given? (Impossibility of defining an inverse of these multiplications.)

2. Give a proof of the rule for differentiating $\mathbf{X}(t)\cdot\mathbf{Y}(t)$.

3. It is given that $\mathbf{X}(t)$ is differentiable and that $|\mathbf{X}(t)| = 1$. Show that $\mathbf{X}(t)$ is orthogonal to $\mathbf{X}'(t)$.

4. Prove that $\mathbf{X}(t) = \mathbf{A}\cos t + \mathbf{B}\sin t$ represents an ellipse. ($\mathbf{A}$ and $\mathbf{B}$ are linearly independent constant vectors.)

5. If $\mathbf{X} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3$ is a unit vector, show that the constants α_i are the direction cosines of the line containing $\mathbf{X}$ which is directed in the same sense as $\mathbf{X}$.

6. Verify the identity of Lagrange:

$$(\mathbf{X} \times \mathbf{Y})\cdot(\mathbf{U} \times \mathbf{V}) = \begin{vmatrix} \mathbf{X}\cdot\mathbf{U} & \mathbf{Y}\cdot\mathbf{U} \\ \mathbf{X}\cdot\mathbf{V} & \mathbf{Y}\cdot\mathbf{V} \end{vmatrix}.$$

7. It is given that $\mathbf{X}(t)$ is differentiable. Show that

$$\mathbf{X}'(t_0) = \lim_{t \to t_0} \frac{\mathbf{X}(t) - \mathbf{X}(t_0)}{t - t_0}$$

by using (1.24).

CHAPTER II

Plane Curves

1. Introduction

From one point of view this chapter could be regarded as unnecessary, since the theory of twisted curves in three-dimensional space is treated in the next chapter, and that theory could be specialized for the case of plane curves. However, there are a number of good reasons for dealing with plane curves separately, quite aside from their specific interest for their own sake. To begin with, plane curves are the simplest objects dealt with in differential geometry, but for all that their study reveals something of the general attitudes and points of view that prevail in differential geometry, even in surface theory. Thus some quite simple developments regarding plane curves foreshadow a good many things to be taken up later in which the circumstances are more complicated. In addition, there are some specific differences in principle worth pointing out with respect to the possible methods of treating plane curves as contrasted with those for space curves.

2. Regular Curves

Experience has shown that it is useful and reasonable to deal in differential geometry (and in other disciplines as well, such as the theory of analytic functions of a complex variable) with a *class* of plane curves called *regular curves*. A regular curve is defined as the locus of points (cf. Fig. 2.1) traced out by the end point of a vector $\mathbf{X}(t)$ in an x_1,x_2-plane:

$$\mathbf{X}(t) = (x_1(t), x_2(t)), \qquad \alpha \le t \le \beta, \tag{2.1}$$

and such that $\mathbf{X}(t)$ satisfies the following conditions:

(a) $\mathbf{X}(t)$ has continuous second derivatives in the interval $\alpha \le t \le \beta$, and
(b) $\mathbf{X}'$, the derivative of $\mathbf{X}(t)$, is nowhere zero.

These conditions merit some discussion. First of all, it might be noted in connection with the condition (a) that for a good deal of the discussion to follow it would be sufficient to require the existence of a continuous first derivative. In general in this book the existence of a certain finite number of derivatives of the functions employed will be assumed, but the minimum number of derivatives needed from case to case will not always be stated. On the other hand, it is *not* desirable to require the functions to be analytic, as is commonly done in the older literature. It is necessary to operate carefully with the tools of analysis, but it is nevertheless geometry rather than analysis that is the subject of this book.

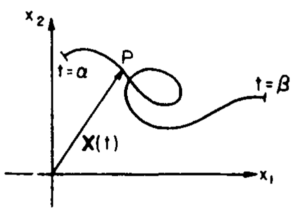

Fig. 2.1 Plane curve defined by a vector.

The condition (b) is more trenchant than condition (a), and it is important to understand why such a condition (which at first sight might seem unnecessarily restrictive) should be imposed. The purpose of the condition—as is also the purpose of analogous conditions for curves and surfaces in three-dimensional space—is *to ensure that the mapping of the t-interval into the x_1,x_2-plane is topological in the small.* By a topological mapping of one object on another is meant here, as in general in mathematics, that the mapping sets up a one-to-one point correspondence that is continuous in both directions. In Fig. 2.1 the curve shown is not the topological image of a t-interval, since it has a double point, and hence it is not in one-to-one correspondence with that interval. However, in the small (that is, in a sufficiently small neighborhood of any point) the mapping is one-to-one if the curve is regular, as is now to be shown. Consider two points $\mathbf{X}(t_0)$, $\mathbf{X}(t_1)$ with $\alpha \leq t_0, t_1 \leq \beta$ and write

$$\begin{aligned}\mathbf{X}(t_1) - \mathbf{X}(t_0) &= (x_1(t_1) - x_1(t_0), x_2(t_1) - x_2(t_0)) \\ &= (x_1'(\xi_1), x_2'(\xi_2))\cdot(t_1 - t_0),\end{aligned}$$

the second line being a result of the mean value theorem; i.e.,

$$x_i(t_0) - x_i(t_1) = x_i'(\xi_i)(t_1 - t_0), \qquad t_0 < \xi_i < t_1.$$

In accordance with some remarks made at the end of the previous chapter, the above equation is put in the form

$$\mathbf{X}(t_1) - \mathbf{X}(t_0) = (t_1 - t_0)\overset{*}{\mathbf{X}}',$$

with $\lim_{t_1 \to t_0} \overset{*}{\mathbf{X}}' = \mathbf{X}'(t_0)$. Since $\mathbf{X}'(t) \neq 0$ anywhere, and this derivative is continuous, it is clear that $\mathbf{X}(t_1) \neq \mathbf{X}(t_0)$ for t_1 near to t_0 but not equal to it, since $\overset{*}{\mathbf{X}}'$ is not the zero vector. In other words, any pair t_0, t_1 of distinct points of the t-interval with $|t_1 - t_0|$ sufficiently small corresponds always to a pair of distinct points in the x_1,x_2-plane. The mapping is therefore locally one-to-one. As a rule in this book, as was stated in the introduction, interest is focused, at least initially, on properties of curves and surfaces in the small, and consequently a double point such as P in Fig. 2.1 is regarded as two distinct points, each of which lies on a different curve segment.

3. Change of Parameters

It is often convenient to shift from one parameter representation of a curve to another. Quite generally, many important results in differential geometry can often be made direct and easy to achieve once a special parametric representation has been tactfully chosen.

In the present case, if $\mathbf{X}(t)$, $\alpha \leq t \leq \beta$, represents a regular curve, the introduction of a new parameter τ, $\gamma \leq \tau \leq \delta$, is brought about by mapping the t-interval in a one-to-one way on a τ-interval by a suitable function $t = \psi(\tau)$. The locus in the plane is of course assumed not to be changed; i.e., $\mathbf{X}(t)$ and $\mathbf{X}(\psi(\tau)) \equiv \mathbf{X}(\tau)$ are assumed to yield the same point for the corresponding values of t and τ. A suitable function $t = \psi(\tau)$ is obtained if it is assumed that

$$\psi'(\tau) \neq 0, \tag{2.2}$$

and in this case the curve will be a regular curve with τ as parameter if $\psi(\tau)$ has a continuous second derivative. This follows because the condition (2.2) means that t is a monotonic function of τ; hence the t-interval and the τ-interval are in one-to-one correspondence. Thus the inverse function $\tau = \phi(t)$ exists and $\phi(t)$ would, like $\psi(t)$, have a continuous second derivative; hence $\mathbf{X}(\psi(\tau))$ has a continuous second derivative and is such that

$$\frac{d\mathbf{X}}{d\tau} = \frac{d\mathbf{X}}{d\psi}\frac{d\psi}{d\tau} \tag{2.3}$$

is not the zero vector since $d\mathbf{X}/d\psi \equiv d\mathbf{X}/dt \neq 0$ holds because the curve was assumed to be regular with t as parameter. (It is convenient on occasion to use the term *regular parameter* in such cases.) Thus the conditions required of $\mathbf{X}(\tau)$ in order that it should represent a regular curve are satisfied.

In particular, it is always possible to introduce as a local parameter one or the other of the coordinates x_1, x_2, for the following reasons. Since $\mathbf{X}'(t) \neq 0$ holds, it follows that $x_1'(t_0)$ and $x_2'(t_0)$ cannot both be zero. Suppose that $x_1'(t_0) \neq 0$. In that case $t = t(x_1)$ is defined as a function of x_1 in a certain neighborhood of the value $x_1(t_0)$ of x_1, hence that

$$\mathbf{X}(x_1) = (x_1, x_2(x_1))$$

is a valid representation of the curve with x_1 as parameter. The curve can then be represented in the more usual way, which dispenses with vectors and defines the curve points by giving one coordinate as a function of the other:

$$x_2 = x_2(x_1). \tag{2.4}$$

This also provides, incidentally, another way of seeing that the condition $\mathbf{X}'(t) \neq 0$ ensures that the mapping into the x_1,x_2-plane is one-to-one in the small: the desired mapping is achieved by a one-to-one orthogonal projection on the x_1-axis.

It is worth pointing out that while the condition (b) is an appropriate sufficient condition for regularity of a curve, it may be violated in spite of the fact that the curve itself has no points that could reasonably be called singular from the geometric point of view. For example, the straight line defined by

$$\mathbf{X}(t) = (t, t),$$

is a regular curve in the sense of the above definition for $-\infty < t < \infty$. The vector

$$\mathbf{X}(t) = (t^3, t^3)$$

clearly yields the same *locus* of points for $-\infty < t < \infty$, but the condition $\mathbf{X}'(t) \neq 0$ is violated for $t = 0$. Evidently, this comes about merely because of an inappropriate choice of the parameter representation. On the other hand, the curve given by

$$\mathbf{X}(t) = (t^2, t^3)$$

has what should properly be called a singularity for $t = 0$ (for reasons that will be pointed out later) in the form of a cusp (cf. Fig. 2.2), which is not due simply to a bad choice of parameter representation. This is a type of ques-

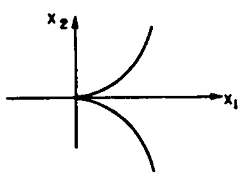

Fig. 2.2 The curve $\mathbf{X} = (t^2, t^3)$.

tion which is considered only in passing in this book. It might be added that such questions, including the classification of types of singularities, can be handled in a rather simple and complete way if the curve is assumed to be analytic (see Pogorelov [P. 5]).

4. Invariance Under Changes of Parameter

It is clear that a regular curve as defined above is a geometrical object which is invariant under transformations of the coordinate axes (for reasons discussed in the preceding chapter). In addition, in differential geometry it is usually regarded as necessary that a geometrical object should be invariant under parameter transformations as well. Here, it is made a matter of definition that two curves $\mathbf{X}(t)$ and $\mathbf{X}(\tau)$ are regarded as the same if $\tau = \phi(t)$ is a transformation of the sort discussed above. However, such an attitude, which means that the plane curves are invariants in this sense under parameter transformations, would be quite inappropriate in mechanics, for example, since *motions* of a particle along a given curve in the time t are not the same if the curve points are traversed at a different rate, although the shape of the trajectory remains the same: different forces are, in fact, required if the motions are different, and hence the two "mechanical objects" consisting of a fixed trajectory and a particle traversing it in different ways are different. Thus it is interesting and important to point out from time to time whether a given entity is, or is not, invariant under parameter transformations.

5. Tangent Lines and Tangent Vectors of a Curve

A tangent line at a point P_0 of a regular curve is commonly defined as a straight line at the point that has as its direction the limiting direction of chords obtained by joining P_0 to points P of the curve near to it and then allowing P to approach P_0. Such chords are evidently line segments joining the end points of vectors $\mathbf{X}(t_0)$ and $\mathbf{X}(t)$, with t_0 and t parameter values corresponding to P_0 and P. The difference quotient $[\mathbf{X}(t) - \mathbf{X}(t_0)]/(t - t_0)$, for $t \neq t_0$, is evidently a vector in a line through the origin that is parallel to the chord joining P_0 and P (see Fig. 2.3). It is assumed that $\mathbf{X}(t)$ is a regular curve; consequently the derivative $\mathbf{X}'(t)$ exists and is given by $\lim_{t\to t_0}[\mathbf{X}(t) - \mathbf{X}(t_0)]/(t - t_0)$. It follows that the vector $\mathbf{X}'(t)$ lies in a line parallel to the limiting direction of the chords under consideration. *The tangent line at point P_0 is now defined as the straight line through P_0 parallel to the direction fixed by the derivative $\mathbf{X}'(t_0)$ of $\mathbf{X}(t)$.* Since $\mathbf{X}'(t) \neq 0$ holds for regular curves,

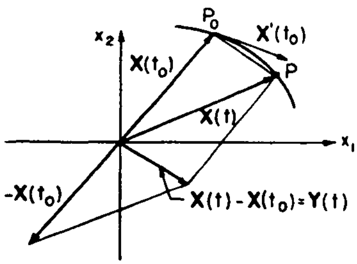

Fig. 2.3 Tangent vector of a curve.

this definition yields a *uniquely determined* line: in fact, one of the main reasons for imposing the condition $\mathbf{X}'(t) \neq 0$ was to achieve that.

The tangent line is a regular curve given by the vector equation

$$\mathbf{T}(r) = \mathbf{X}(t_0) + r\mathbf{X}'(t_0), \qquad -\infty < r < \infty, \tag{2.5}$$

in which (cf. Fig. 2.4) the point of tangency is the point fixed by $t = t_0$, and r is the parameter on the tangent line. [Problem 1 at the end of the chapter requires a proof that $\mathbf{T}(r)$ represents a regular curve.]

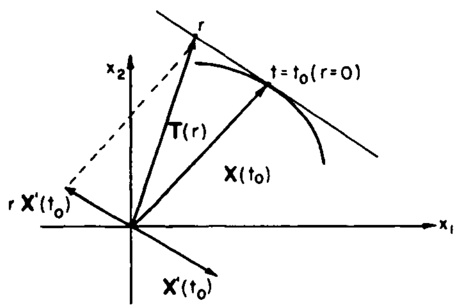

Fig. 2.4 Tangent line of a curve.

In the differential geometry of curves it is advantageous to define tangent *vectors* localized at the points of the curves. The derivative $\mathbf{X}'(t)$ could be, and by some writers is, defined as a tangent vector, but this definition would have the disadvantage of placing all of the tangent vectors at the origin rather than at the appropriate points on the curve. A reasonable way out is to use the possibility afforded by Euclidean geometry (but which is not available in other geometries to be studied later in this book) of moving vectors parallel to themselves. Thus the tangent vector at a point $\mathbf{X}(t_0)$ of the curve is *defined* as *the vector obtained by translating* $\mathbf{X}'(t_0)$ *parallel to itself to this point.* It is then still denoted by $\mathbf{X}'(t_0)$. It would also be possible to proceed in another way to achieve the same end result by first translating the origin of the coordinate system in the plane to a point on the curve and then taking the derivative $\mathbf{X}'(t)$. It is clear, and in any case easy to show, that both procedures would lead to the same tangent vector.[1]

It is to be noted that the tangent vector $\mathbf{X}'(t)$ is invariant under coordinate transformations but not under parameter transformations, since $d\mathbf{X}/d\tau = (d\mathbf{X}/dt)(dt/d\tau)$. However, the notion of a *tangent line* is seen to be independent of the choice of a parameter representation of $\mathbf{X}(t)$.

6. Orientation of a Curve

The choice of a particular parameter representation fixes a direction of travel along the curve, which in turn has a relation to the tangent vector defined above. To study this matter, the vector $\mathbf{Y}(t)$ is defined as follows:

$$\mathbf{Y}(t) = \mathbf{X}(t) - \mathbf{X}(t_0) = (t - t_0)\overset{*}{\mathbf{X}}', \tag{2.6}$$

with $\overset{*}{\mathbf{X}}'$ a vector that tends to the derivative $\mathbf{X}'(t_0)$ as $t \to t_0$ (again a derivative with a * is used). The vector $\mathbf{Y}$ is parallel to the secant $\overline{P_0P}$, as shown in Fig. 2.3. Hence, if $|t - t_0|$ is small enough, $\mathbf{Y}(t)$ has nearly the direction of the tangent vector $\mathbf{X}'(t_0)$ if $(t - t_0)$ is positive and nearly the opposite direction if $(t - t_0)$ is negative. In other words, $\mathbf{X}'(t_0)$ *points into that half-plane bounded by a line normal to* $\mathbf{X}'(t_0)$ *in which the curve points lie for* $t > t_0$, *i.e., for increasing values of* t. It is thus reasonable to say that the direction of the tangent vector fixes the direction of travel along the curve when the parameter values increase. If the parameter transformation from t into $-t$ is made, it is clear that the direction of $\mathbf{X}'(t)$ is reversed; in fact, that is the case for any transformation $\tau = \tau(t)$ if τ' is negative, as can be seen from (2.3). The orientation of a curve is thus not in general an invariant property with respect to all parameter transformations, but only for those that satisfy the inequality $dt/d\tau > 0$.

One of the reasons for calling the point $t = 0$ of the curve $\mathbf{X}(t) = (t^2, t^3)$, $-\infty < t < \infty$ (see Fig. 2.2) a singular point can now be given. Since t changes sign on passing through $t = 0$, an abrupt reversal in the direction of travel takes place, although t changes in a monotonic way. A still better reason, perhaps, for calling the point a singularity is that the limiting direction of chords could be any direction if the chords were to be drawn through

[1] Some stress is put on this matter because it is often a sore point in mechanics, and much confusion can arise there because of a failure to recognize that it is only very exceptionally that vectors can be moved along their lines, and still less parallel to themselves, while continuing to represent a given physical entity correctly.