Elementary Differential Geometry

Barrett O'Neill

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA

ACADEMIC PRESS New York and London

ACADEMIC PRESS INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London W.1

Library of Congress Catalog Card Number: 66-14468

PRINTED IN THE UNITED STATES OF AMERICA

Preface

This book is an elementary account of the geometry of curves and surfaces. It is written for students who have completed standard first courses in calculus and linear algebra, and its aim is to introduce some of the main ideas of differential geometry.

The traditional undergraduate course in differential geometry has changed very little in the last few decades. By contrast, geometry has been advancing very rapidly at the research level, and there is general agreement that the undergraduate course needs to be brought up to date. I have tried to think through the classical material, to prune and augment it, and to write down the results in a reasonably clean and modern mathematical style. However, I have used a new idea only if it really pays its way by simplifying and clarifying the exposition.

Chapter I establishes the language of the book—a language compounded of familiar parts of calculus and linear algebra. Chapter II describes the method of "moving frames," which is introduced, as in elementary calculus, to study curves in space. In Chapter III we investigate the rigid motions of space, in terms of which congruence of curves (or surfaces) in space is defined in the same fashion as congruence of triangles in the plane.

Chapter IV requires special comment. The main weakness of classical differential geometry was its lack of any adequate definition of *surface*. In this chapter we decide just what a surface is, and show that each surface has a differential and integral calculus of its own, strictly comparable with the familiar calculus of the plane. This exposition provides an introduction to the notion of *differentiable manifold*, which has become indispensable to those branches of mathematics and its applications based on the calculus.

The next two chapters are devoted to the geometry of surfaces in 3-space. Chapter V stresses intuitive and computational aspects to give geometrical meaning to the theory presented in Chapter VI. In the final chapter, although our methods are unchanged, there is a radical shift of viewpoint. Roughly speaking, we study the geometry of a surface *as seen by its inhabi-*

tants, with no assumption that the surface is to be found in ordinary three-dimensional space.

No branch of mathematics makes a more direct appeal to the intuition than geometry. I have sought to emphasize this by a large number of illustrations, which form an integral part of the text. A set of exercises appears at the end of each section; these range from routine tests of comprehension to more seriously challenging problems.

In teaching from preliminary versions of this book, I have usually covered the background material in Chapter I rather rapidly, and have not devoted any classroom time to Chapter III (hence also Section 8 of Chapter VI). A course in the geometry of curves and surfaces in space might consist of: Chapter II, Chapter IV (omit Sections 6 and 8), Chapter V, and Chapter VI (omit Sections 6 and 7). This is essentially the content of the traditional undergraduate course in differential geometry, with clarification of the concepts of surface and mapping of surfaces.

The omitted sections in the list above are used only in Chapter VII. This final chapter, an extensive account of two-dimensional Riemannian geometry, is in a sense the goal of the book. Rather than shift the discourse to higher dimensions, I have preferred to retain dimension 2, so that this more sophisticated view of geometry will develop directly from the special case of surfaces in 3-space. Chapter VII is long, and on a first reading Theorem 5.9 and Sections 6 and 7 may well be omitted. Serious use of differential equations theory has been largely avoided in the early chapters; however, some acquaintance with the fundamentals of the subject will be helpful in Chapter VII.

Los Angeles, California B. O'N.

Introduction

This book presupposes a reasonable knowledge of elementary calculus and linear algebra. It is a working knowledge of the fundamentals that is actually required. The reader will, for example, frequently be called upon to *use* the chain rule for differentiation, but its proof need not concern us.

Calculus deals mostly with real-valued functions of one or more variables, linear algebra with functions (linear transformations) from one vector space to another. We shall need functions of these and other types, so we give here general definitions which cover all types.

A *set* S is a collection of objects which are called the *elements* of S. A set A is a *subset* of S provided each element of A is also an element of S.

A *function* f from a set D to a set R is a rule that assigns to each element x of D a unique element $f(x)$ of R. The element $f(x)$ is called the *value* of f at x. The set D is called the *domain* of f; the set R is often called the *range* of f. If we wish to emphasize the domain and range of a function f, the notation $f\colon D \rightarrow R$ is used. Note that the function is denoted by a single letter, say f, while $f(x)$ is merely a value of f.

Many different terms are used for functions—mappings, transformations, correspondences, operators, and so on. A function can be described in various ways, the simplest case being an explicit formula such as

$$f(x) = 3x^2 + 1,$$

which we may also write as $x \rightarrow 3x^2 + 1$.

If both f_1 and f_2 are functions from D to R, then $f_1 = f_2$ means that $f_1(x) = f_2(x)$ for all x in D. This is not a definition, but a logical consequence of the definition of function.

Let $f\colon D \rightarrow R$ and $g\colon E \rightarrow S$ be functions. In general, the *image* of f is the subset of R consisting of all elements of the form $f(x)$; it is usually denoted by $f(D)$. Now if this image also happens to be a subset of the domain E of g, it is possible to combine these two functions to obtain the *composite function* $g(f)\colon D \rightarrow S$. By definition, $g(f)$ is the function whose value on each element x of D is the element $g(f(x))$ of S.

If $f\colon D \to R$ is a function and A is a subset of D, then the *restriction* of f to A is the function $f \mid A\colon A \to R$ defined by the same rule as f, but applied only to elements of A. This seems a rather minor change, but the function $f \mid A$ may have properties quite different from f itself.

Here are two vital properties which a function may possess. A function $f\colon D \to R$ is *one-to-one*, provided that if x and y are any elements of D such that $x \neq y$, then $f(x) \neq f(y)$. A function $f\colon D \to R$ is *onto* (or *carries* D *onto* R) provided that for every element y of R there is at least one element x of D such that $f(x) = y$. In short, the image of f is the entire set R. For example, consider the following functions, each of which has the real numbers as both domain and range:

(1) The function $x \to x^3$ is both one-to-one and onto.
(2) The exponential function $x \to e^x$ is one-to-one, but not onto.
(3) The function $x \to x^3 + x^2$ is onto, but not one-to-one.
(4) The sine function $x \to \sin x$ is neither one-to-one nor onto.

If a function $f\colon D \to R$ is both one-to-one and onto, then for each element y of R there is one and only one element x such that $f(x) = y$. By defining $f^{-1}(y) = x$ for all x and y so related, we obtain a function $f^{-1}\colon R \to D$ called the *inverse* of f. Note that the function f^{-1} is also one-to-one and onto, and that *its* inverse function is the original function f.

Here is a short list of the main notations used throughout the book, in order of their appearance in Chapter I:

$\mathbf{p}, \mathbf{q}, \ldots$	points	(Sec. 1)
$f, g, \ldots$	real-valued functions	(Sec. 1)
$\mathbf{v}, \mathbf{w}, \ldots$	tangent vectors	(Sec. 2)
$V, W, \ldots$	vector fields	(Sec. 2)
$\alpha, \beta, \ldots$	curves	(Sec. 4)
$\phi, \psi, \ldots$	differential forms	(Sec. 5)
$F, G, \ldots$	mappings	(Sec. 7)

In Chapter I we define these concepts for Euclidean 3-space. (Extension to arbitrary dimensions is virtually automatic.) In Chapter IV we show how these concepts can be adapted to a surface.

A few references are given to the brief bibliography at the end of the book; these are indicated by numbers in square brackets.

CHAPTER I

Calculus on Euclidean Space

As mentioned in the Preface, the purpose of this initial chapter is to establish the mathematical language used throughout the book. Much of what we do is simply a review of that part of elementary calculus dealing with differentiation of functions of three variables, and with curves in space. Our definitions have been formulated so that they will apply smoothly to the later study of surfaces.

1 Euclidean Space

Three-dimensional space is often used in mathematics without being formally defined. It is said to be the space of ordinary experience. Looking at the corner of a room, one can picture the familiar process by which rectangular coordinate axes are introduced and three numbers are measured to describe the position of each point. A precise definition which realizes this intuitive picture may be obtained by this device: instead of saying that three numbers *describe the position* of a point, we define them to *be* a point.

1.1 Definition *Euclidean 3-space* $\mathbf{E}^3$ is the set of all ordered triples of real numbers. Such a triple $\mathbf{p} = (p_1, p_2, p_3)$ is called a *point* of $\mathbf{E}^3$.

In linear algebra, it is shown that $\mathbf{E}^3$ is, in a natural way, a vector space over the real numbers. In fact, if $\mathbf{p} = (p_1, p_2, p_3)$ and $\mathbf{q} = (q_1, q_2, q_3)$ are points of $\mathbf{E}^3$, their *sum* is the point

$$\mathbf{p} + \mathbf{q} = (p_1 + q_1, p_2 + q_2, p_3 + q_3).$$

The *scalar product* of a point $\mathbf{p} = (p_1, p_2, p_3)$ by a number a is the point

$$a\mathbf{p} = (ap_1, ap_2, ap_3).$$