CALCULUS OF VARIATIONS

By

GILBERT AMES BLISS
Professor of Mathematics, The University of Chicago

Published for
THE MATHEMATICAL ASSOCIATION OF AMERICA
by
THE OPEN COURT PUBLISHING COMPANY
CHICAGO · ILLINOIS

PREFACE

This book is the first of a series of monographs on mathematical subjects which are to be published under the auspices of the Mathematical Association of America and whose publication has been made possible by a very generous gift to the Association by Mrs. Mary Hegeler Carus as trustee for the Edward C. Hegeler Trust Fund. The purpose of the monographs is to make the essential features of various mathematical theories accessible and attractive to as many persons as possible who have an interest in mathematics but who may not be specialists in the particular theory presented, a purpose which Mrs. Carus has very appropriately described to be "the diffusion of mathematical and formal thought as contributory to exact knowledge and clear thinking, not only for mathematicians and teachers of mathematics but also for other scientists and the public at large."

The attainment of this end will not always be easy for authors who have long specialized in unraveling the intricacies of the domains in which their principal activities lie, and the clientele of readers which they may reasonably hope to interest will vary greatly with the subjects presented. It would obviously be unwise to regard this first attempt as in any final sense a model for the many monographs which it is hoped will follow. Later authors will doubtless profit much by the experiences of those who have written before, but varieties of subjects and types of readers to be addressed are likely to require an equally large variety of methods of presentation. It is possible that some monographs will be entirely descriptive or historical in character, others devoted to the treatment in detail of special mathematical questions which can be approached without elaborate prerequisite study, and still others of types not yet devised but which are certain to be suggested as the series progresses. One can readily foresee the beneficial influence which the monographs will have in encouraging and developing types of descriptive mathematical writing suited to the very laudable purposes for which the series has been inaugurated.

The theory to which the present monograph is devoted, the calculus of variations, is one whose development from the beginning has been interlaced with that of the differential and integral calculus. Without any knowledge of the calculus one can readily understand at least the geometrical or mechanical statements of many of the problems of the calculus of variations and the character of their solutions, as an examination of the chapters to follow will show. Thus if two points not in the same vertical line are given we may ask for the curve joining them down which a marble starting with a given initial velocity will roll from one point to the other in the shortest time. The solution is a piece of an inverted cycloid, and a cycloid is the curve described by a point on the rim of a wheel as the wheel rolls along the ground. Or if two points above a horizontal x-axis are given we may seek to find the curve which joins them and which when rotated around the x-axis generates a surface of revolution of minimum area. The solution curve will have sometimes one and sometimes the other of two forms. The first of these is the broken line consisting of the two perpendiculars from the points to the x-axis and the portion of the axis between them, in which case the minimum surface consists of two circular disks. The second is an arc of a catenary, and the form of a catenary is that which a chain naturally takes when suspended from two pegs. The surface generated in this latter case is the capstan-shaped surface assumed by a soap film suspended between two wire circles having a common axis.

The discovery and justification of the results which have just been described, apart from their simple statement, do require, however, acquaintance with the principles of the calculus, and in the following pages it is assumed that the reader has such an acquaintance. This should not deter others who may be interested from examining the introductions to the various chapters and the italicized theorems throughout the book, many of which should be perfectly intelligible to everyone. The only place where results not usually deduced in the ordinary calculus course are used is in the last chapter, where some properties of differential equations are required which have already been clearly illustrated in the three preceding chapters, and which are described in detail in the text.

In selecting material for presentation it seemed desirable to begin by studying special problems rather than the general theory. The first chapter of the book describes the historical setting out of which the theory of the calculus of variations grew, and the character of some of the simpler problems. The next three chapters are devoted to the development in detail of the known results for three special problems which illustrate in excellent fashion the essential characteristics of the general theory contained in Chapter V with which the book concludes. The author was influenced in this selection by several considerations. In the first place the theory of the special problems here presented requires only analysis of a concrete sort in which one is much aided by intuition while accumulating experiences which assist effectively in understanding later the notions of the general theory. In the second place the theory of these problems, though well known, is scattered in various places in treatises and memoirs on the calculus of variations, and the presentation of it in collected form should therefore be useful as well as instructive. Finally it is a fact that the modern theory of the calculus of variations has been presented for the most part in elaborate mathematical treatises and is not readily accessible except to the specialist. The elementary discussions of the theory, in the larger more general treatises on analysis and also in separate form, usually lay their emphasis upon the deduction of the differential equations of the minimizing curves for various types of problems, and relatively little upon other aspects of the theory. It is doubtless partly for this reason that in applied mathematics much more use has hitherto been made of these differential equations and their solutions than of the further properties of minimizing curves, though it is well known that in many cases these further properties are closely related to interesting conditions for stability in associated problems of mechanics.

Such are the reasons why it seemed desirable to the author to present in this book the theory of special problems with some completeness, even if limitations of space should permit only a few of them to be discussed. It must be admitted that in the literature of the calculus of variations there are not many particular cases to which the general theory has been thoroughly applied. The assembling of as many such problems as possible and the completion of others would be a work of great usefulness and interest.

At the end of Chapter V is a list of the books on the calculus of variations with a few other references of importance for the topics considered in the text. The notes, indicated serially in the text by superscripts, follow this list of references.

G. A. Bliss

The University of Chicago
October, 1924

TABLE OF CONTENTS

CHAPTER I

TYPICAL PROBLEMS OF THE CALCULUS OF VARIATIONS

1. *The invention of the calculus.* When the student of mathematics pauses to look back upon the achievements of mathematicians of the past he must be impressed with the fact that the seventeenth century was a most important epoch in the development of modern mathematical analysis, since to the mathematicians of that period we owe the invention of the differential and integral calculus. At first the calculus theory, if indeed at that time it could be called such, consisted of isolated and somewhat crude methods of solving special problems. In the domain of what we now call the integral calculus, for example, an Italian mathematician named Cavalieri (1598–1647) devised early in the seventeenth century a summation process, called the method of indivisibles, by means of which he was able to calculate correctly many areas and volumes. His justification of his device was so incomplete logically, however, that even in those relatively uncritical times his contemporaries were doubtful and dissatisfied. Somewhat later two French mathematicians, Roberval (1602–75) and Pascal (1623–62), and the Englishman Wallis (1616–1703), improved the method and made it more like the summation processes of the integral calculus of today. In the case of the differential calculus we find that before the final quarter of the seventeenth century Descartes (1596–1650), Roberval, and Fermat (1601–65) in France, and Barrow

(1630–77) in England, all had methods of constructing tangents to curves which were pointing the way toward the solution of the fundamental problem of the differential calculus as we formulate it today, namely, that of determining the slope of the tangent at a point of a curve.

At this important stage there appeared upon the scene two men of extraordinary mathematical insight, Newton (1642–1727) in England, and Leibniz (1646–1716) in Germany, who from two somewhat different standpoints carried forward the theory and applications of the calculus with great strides. It is a mistake, though we often find it an easy convenience, to regard these two great thinkers as having invented the calculus out of a clear sky. Newton was in fact a close student of the work of Wallis, and a pupil of Barrow whom he succeeded as professor of mathematics at Cambridge, while we know that Leibniz visited Paris and London early in his career and that he there became acquainted with the most advanced mathematics of his day. No one could successfully contest the fact, however, that these two men were the most able spokesmen and investigators of the seventeenth-century school of mathematicians to which we owe the gradual evolution of the calculus.

In spite of the great abilities of Newton and Leibniz the underlying principles of the calculus as exposed by them seem to us from our modern viewpoint, as indeed to their contemporaries and immediate successors, somewhat vague and confusing. The difficulty lies in the lack of clearness at that early time, and for more than a century thereafter, in the conceptions of infinitesimals and limits upon which the calculus rests, a difficulty which has been overcome only by the systematic study of the theory of limits inaugurated by Cauchy (1789–1857) and continued by Weierstrass (1815–97), Riemann (1826–66), and many others.

2. *Maxima and minima.* Among the earliest problems which attracted the attention of students of the calculus were those which require the determination of a maximum or a minimum. Fermat had devised as early as 1629 a procedure applicable to such problems, depending upon principles which in essence, though not in notation, were those of the modern differential calculus. Somewhat nearer to the type of reasoning now in common use are the methods which Newton and Leibniz applied to the determination of maxima and minima, methods which are also characteristic of their two conceptions of the fundamental principles of the differential calculus. Newton argued, in a paper written in 1671 but first published in 1736, that a variable is increasing when its rate of change is positive, and decreasing when its rate is negative, so that at a maximum or a minimum the rate must be zero. Leibniz, on the other hand, in a paper which he published in 1684, conceived the problem geometrically. At a maximum or a minimum point of a curve the tangent must be horizontal and the slope of the tangent zero.

At the present time we know well that from a purely analytical standpoint these two methods are identical. The derivative

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \qquad (1)$$

of a function $f(x)$ represents both the rate of change of $f(x)$ with respect to x and the slope of the tangent at a point on the graph of $f(x)$. For in the first place the fraction in the second member of equation (1) is the average rate of change of $f(x)$ with respect to x on the interval from x to $x+\Delta x$, and its limit as the interval is shortened is therefore rightly called the rate of change of $f(x)$ at the initial value x of the interval. In the second place this same quotient is the slope of the secant PQ in Figure 1, and its limit is the slope of the tangent at P. Thus by the reasoning of either Newton or Leibnitz we know that the maxima and minima of $f(x)$ occur at the values of x where the derivative $f'(x)$ is zero.

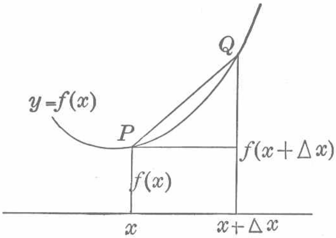

FIG. 1

It was not easy for the seventeenth-century mathematician to deduce this simple criterion that the derivative $f'(x)$ must vanish at a maximum or a minimum of $f(x)$. He was immersed in the study of special problems rather than general theories, and had no well-established limiting processes or calculus notations to assist him. It was still more difficult for him to advance one step farther to the realization of the significance of the second derivative $f''(x)$ in distinguishing between maximum and minimum values. Leibniz in his paper of 1684 was the first to give the criterion. In present-day parlance we say that $f'(a)=0$, $f''(a)\geqq 0$ are necessary conditions for the value $f(a)$ to be a minimum, while the conditions $f'(a)=0$, $f''(a)>0$ are sufficient to insure a minimum. For a maximum the inequality signs must be changed in sense.

It will be noted that the conditions just stated as necessary for a minimum are not identical with those which are sufficient. We shall see in Chapter V that a similar undesirable and much more baffling discrepancy occurs in the calculus of variations. For the simple problem of minimizing a function $f(x)$ the doubtful intermediate case when $f'(a)$ and $f''(a)$ are both zero was discussed by Maclaurin (1698–1746) who showed how higher derivatives may be used to obtain criteria which are both necessary and sufficient. For the calculus of variations the corresponding problem offers great difficulty and has never been completely solved.

3. *Two problems of the calculus of variations which may be simply formulated.* When one realizes the difficulty with which the late seventeenth-century school of mathematicians established the first fundamental principles of the calculus and their applications to such elementary problems in maxima and minima as the one which has just been described, it is remarkable that they should have conceived or attempted to solve with their relatively crude analytical machinery the far more difficult maximum and minimum problems of the calculus of variations which were at first proposed. It is an interesting fact that these early problems were not by any means the least complicated ones of the calculus of variations, and we shall do well therefore to introduce ourselves to the subject by looking first at two others which are easier to describe to one who has not already amused himself by browsing in this domain of mathematics.

The simplest of all the problems of the calculus of variations is doubtless that of determining the shortest arc joining two given points. The co-ordinates of these points will always be denoted by (x_1, y_1) and (x_2, y_2) and we may designate the points themselves when convenient simply by the numerals 1 and 2. If the equation of an arc is taken in the form

$$y=y(x) \qquad (x_1 \leqq x \leqq x_2)$$

then the conditions that it shall pass through the two given points are

$$(2) \qquad y(x_1)=y_1 , \qquad y(x_2)=y_2 ,$$

and we know from the calculus that the length of the arc is given by the integral

$$I=\int_{x_1}^{x_2} \sqrt{1+y'^2}\, dx ,$$

where in the evaluation of the integral y' is to be replaced by the derivative $y'(x)$ of the function $y(x)$ defining the arc. There is an infinity of curves $y=y(x)$ joining the points 1 and 2. The problem of finding the shortest one is equivalent analytically to that of finding in the class of functions $y(x)$ satisfying the conditions (2) one which makes the integral I a minimum.

In the more elementary minimum problem of Section 2 above a function $f(x)$ is given and a value $x=a$ is sought for which the corresponding value $f(a)$ is a minimum. In the shortest-distance problem the integral I takes the place of $f(x)$, and instead of a value $x=a$ making $f(a)$ a minimum we seek to find an arc E_{12} joining the points 1 and 2 which shall minimize I. The analogy between the two problems is more perspicuous if we think of the length I as a function $I(E_{12})$ whose value is uniquely determined when the arc E_{12} is given, just as $f(x)$ in the former case was a function of the variable x.

There is a second problem of the calculus of variations, of a geometrical-mechanical type, which the principles of the calculus readily enable us to express also in analytic form. When a wire circle is dipped in a soap solution and withdrawn, a circular disk of soap film bounded by the circle is formed. If a second smaller circle is made to touch this disk and then moved away the two circles will be joined by a surface of film which is a surface of revolution in the particular case when the circles are parallel and have their centers on the same axis perpendicular to their planes. The form of this surface is shown in Figure 2. It is provable by the principles of mechanics, as one may readily surmise intuitively from the elastic properties of a soap film, that the surface of revolution so formed must be one of minimum area, and the problem of determining the shape of the film is equivalent therefore to that of determining such a minimum surface of revolution passing through two circles whose relative positions are supposed to be given as indicated in the figure.

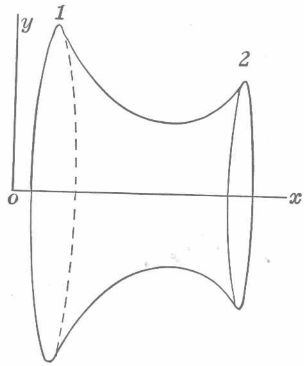

FIG. 2

In order to phrase this problem analytically let the common axis of the two circles be taken as the x-axis, and let the points where the circles intersect an xy-plane through that axis be 1 and 2. If the meridian curve of the surface in the xy-plane has an equation $y=y(x)$ then the calculus formula for the area of the surface is 2π times the value of the integral

$$I=\int_{x_1}^{x_2} y\sqrt{1+y'^2}\,dx\,.$$

The problem of determining the form of the soap film surface between the two circles is analytically that of finding in the class of arcs $y=y(x)$ whose ends are at the points 1 and 2 one which minimizes the last-written integral I.

4. *The problem of Newton.* It was remarked above that the earliest problems of the calculus of variations were not by any means the simplest. In his *Principia* (1686)[1] Newton states without proof certain conditions which must be satisfied by a surface of revolution which is so formed that it will encounter a minimum resistance when moved in the direction of its axis through a resisting medium. A particular case of the problem of finding such a surface is the well-known one of determining the form of a projectile which for a specified initial velocity will give the longest range. In practical ballistics it turns out that one of the most difficult parts of the investigation of this question lies in the experimental determination of the retardation law for bodies moving in the air at high rates of speed. Newton assumed a relatively simple law of resistance for bodies moving in a resisting medium which does not agree well with our experience with bodies moving in the air, but on the basis of which he was able to find a condition characterizing the meridian curves of the surfaces of revolution which encounter minimum