LECTURES ON APPLIED MATHEMATICS

VOLUME 2

The Calculus of Variations

by Francis D. Murnaghan

THE CALCULUS OF VARIATIONS

by ***FRANCIS D. MURNAGHAN***

Consultant, Applied Mathematics Laboratory
The David Taylor Model Basin

SPARTAN BOOKS · WASHINGTON, D. C. 1962

Preface

This book is based on lectures given at the Applied Mathematics Laboratory of the David Taylor Model Basin. It is devoted to the calculus of variations, with special attention to its applications in mechanics and wave propagation. Our aim has been to give more than a superficial account of the subject, treating only extremal or stationary curves. In order to keep this treatment self-contained and free of formidable mathematical difficulties, we have made the necessary differentiability assumptions which are usually satisfied in the applications of the calculus of variations to problems of nature.

We have emphasized the parametric treatment, since this method simplifies the passage from the Lagrangian to the Hamiltonian point of view, or vice versa.

The distinction between extremal and minimal curves is clearly explained, and Legendre's condition for a minimal curve is proved. Extremal fields and the Hilbert Invariant Integral are covered in some detail, the essential distinction between plane problems and problems in multidimensional space, so far as extremal fields are concerned, being carefully explained.

The Weierstrass E-function is treated, and the connection between the calculus of variations and Rayleigh quotients and the method of Rayleigh-Ritz—so important in vibration problems—is clearly shown. The analysis of wave propagation gives the analogue for earthquake waves of Snell's law of refraction. The book closes with a short account of multiple integral problems and a discussion of the useful maximum-minimum principle, which we owe to Courant.

Care has been taken to make the book self-contained, and details of the proofs of the basic mathematical theorems are provided.

Washington, D. C.
May, 1962

Francis D. Murnaghan

Contents

1

The Lagrangian Function and the Parametric Integrand

The problems of the calculus of variations which we shall treat in these pages belong to one or the other of two types. The simplest example of the first of these two types may be stated as follows:

Given two points in a plane, or in 3-dimensional Euclidean space, does there exist a curve of shortest length connecting these two points, and if so, is this curve unambiguously determinate? We shall refer to the problems of the calculus of variations which belong to the type to which this problem belongs as problems of Type 1. An example of the second of the two types may be stated as follows:

Given a closed curve in 3-dimensional Euclidean space, does there exist in this space, a surface, having this curve as its boundary, whose area is least, and if so, is this surface unambiguously determinate? We shall refer to the problems of the calculus of variations which belong to the type to which this problem belongs as problems of Type 2.

Problems of the calculus of variations of Type 1 are *curve* problems, in which the curves we are concerned with may be plane curves or space curves or, indeed, curves in a space of any number of dimensions. The problem of the motion of a mechanical system may be conveniently stated as a problem of Type 1 of the calculus of variations. If the mechanical system has n degrees of freedom, $n = 1, 2, \cdots$, we write its generalized coordinates as a column, or $n \times 1$, matrix x, so that the jth coordinate, $j = 1, \cdots, n$, is denoted by x^j. The velocity $n \times 1$ matrix x_t is the derivative of x with respect to the time t so that the jth element of x_t is x_t^j, $j = 1, \cdots,$

n, and the Lagrangian function $L = T - V$ of the system is a function of the two $n \times 1$ matrices x and x_t, and also, if the mechanical system is nonconservative, of the time t. Here, T is the kinetic energy and V is the potential energy of the mechanical system. We assume that L is a continuous function of the matrices x, x_t, and also, if the system is nonconservative, of t; by the statement that L is a continuous function of the two matrices x and x_t we mean that L is a continuous function of the $2n$ elements x^j, x_t^k, $j = 1, \cdots, n$, $k = 1, \cdots, n$, of these matrices. If, then, $x = x(t), a \leqslant t \leqslant b$, is any smooth curve C in the n-dimensional coordinate space of the mechanical system, so that x and x_t are, along C, continuous functions of t over the interval $a \leqslant t \leqslant b$, the Lagrangian function L is, along C, a continuous function of t over the interval $a \leqslant t \leqslant b$, and we may consider the integral

$$I = \int_a^b L\,dt$$

This integral is a number whose value depends on and is unambiguously determined by the curve C, just as the length of a rectifiable curve in a plane, or in 3-dimensional Euclidean space, depends on and is unambiguously determined by the curve. The motion of the mechanical system is such that its paths, i.e., the curves $x = x(t)$, $a \leqslant t \leqslant b$, which describe this motion, are such that the integral $I = \int_a^b L\,dt$ has, when evaluated along any one of these paths, a stationary value (without being, necessarily, a minimum or a maximum). Thus the integral I plays, for the mechanical system, the role played by the length of a rectifiable curve in the introductory example we have given of problems of Type 1 of the calculus of variations, but there is one essential difference: I has, when evaluated along a path of the mechanical system, merely a stationary value while the length integral is actually a minimum when the curve of integration gives it a stationary value. In this connection we observe that the time-coordinate space of a mechanical system is what is called a numerical, rather than a metrical, space; the concept of distance between two of its points is not defined, so that it does not make sense to speak of nearby points of this coordinate space nor of the length of a curve in it. We may, however, endow this coordinate space with a metric and we shall do this by assigning to it the ordinary Euclidean metric in accordance with which the distance between any two points $\begin{vmatrix} t \\ x \end{vmatrix}$ and $\begin{vmatrix} t' \\ x' \end{vmatrix}$ is the magnitude of the matrix $\begin{vmatrix} t' - t \\ x' - x \end{vmatrix}$.

In order to treat most simply the problem of Type 1 of the calculus of variations which is furnished by the motion of a mechanical system it is convenient (particularly when the system is nonconservative so that the

Lagrangian function L involves explicitly not only the two $n \times 1$ matrices x and x_t but also the time variable t) to place this variable on an equal footing with the n elements of the coordinate matrix x. To do this we replace our n-dimensional coordinate space by a $(n + 1)$-dimensional time-coordinate space. We denote $n + 1$ by N and introduce the $N \times 1$ matrix $X = \begin{vmatrix} t \\ x \end{vmatrix}$ whose first element is t and whose remaining n elements are those of the $n \times 1$ matrix x. A smooth curve in the N-dimensional time-coordinate space is furnished by a formula $X = X(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, where τ is any convenient independent variable, or parameter, of which the $N \times 1$ matrix $X(\tau)$ is a continuously differentiable function. If we choose τ to be t itself, the first of the N equations implied by the formula $X = X(\tau)$ is simply $t = \tau$ but, in general, this equation will be replaced by $t = t(\tau)$, where $t(\tau)$ is either smooth, i.e., possesses a continuous derivative, or is at least piecewise-smooth over the interval $\alpha \leqslant \tau \leqslant \beta$. By the words piecewise-smooth, we mean that $t(\tau)$, while continuous over $\alpha \leqslant \tau \leqslant \beta$, may fail to be differentiable at a finite number of interior points of this interval; at each of these points it possesses a right-hand and a left-hand derivative but these derivatives are not equal. At all points where t_τ is defined it is, by hypothesis, continuous. We make one further restriction on the function $t = t(\tau)$; namely, we assume that $t_\tau > 0$ at all points where t_τ is defined, and this implies that at each of the finite number of points at which t_τ is possibly undefined the right-hand and left-hand derivatives of t with respect to τ are nonnegative since the right-hand derivative, for example, of t at $\tau = \tau_1$, say, is the limit, as $\delta \to 0$ through positive values, of $t_\tau(\tau_1 + \delta)$. The reason for this restriction is as follows: We regard t, which was the parameter or independent variable used by Lagrange, as a master, or control, parameter and we do not wish any other parameter τ to sometimes increase and sometimes decrease as t increases. Thus we do not wish t_τ to change sign over the interval $\alpha \leqslant \tau \leqslant \beta$. We could satisfy this wish by requiring that $t_\tau \leqslant 0$ instead of $t_\tau > 0$, but, since a mere change of sign of τ changes the inequality $t_\tau < 0$ into $t_\tau > 0$ there is no real loss of generality in requiring that t_τ be >0 at all the points of the interval $\alpha \leqslant \tau \leqslant \beta$ at which it is defined. We assume, further, that the number of points, if any, of the interval $\alpha \leqslant \tau \leqslant \beta$ at which $t_\tau = 0$ is finite.

When we pass from the master parameter t to any other allowable parameter τ by means of a formula $t = t(\tau)$, $\alpha \leqslant \tau \leqslant \beta$, the integrand of the integral I is changed from the Lagrangian function L to the product of L by t_τ:

$$I = \int_a^b L\,dt = \int_\alpha^\beta L t_\tau\,d\tau; \qquad a = t(\alpha), \quad b = t(\beta)$$

We denote this new integrand by F and we refer to F as the parametric integrand, the original integrand L being termed the nonparametric, or Lagrangian, integrand of our problem of Type 1 of the calculus of variations. Under an allowable change of parameter $\tau \to \tau'$ furnished by a formula $\tau = \tau(\tau')$, $\alpha' \leqslant \tau' \leqslant \beta'$, L remains unaffected while $t_\tau \to t_{\tau'} = t_\tau \tau_{\tau'}$, where $\tau_{\tau'}$ is positive save, possibly, at a finite number of points of the interval $\alpha' \leqslant \tau' \leqslant \beta'$. Since the value of an integral is insensitive to changes of the integrand at a finite number of points, we shall refer to $\tau_{\tau'}$ as positive (rather than nonnegative). Then the transformation $\tau \to \tau'$ of the independent variable, or parameter, induces the transformation $F \to F' = F\tau_{\tau'}$ of the integrand of our integral I. This change of the integrand is necessary to ensure the invariance, or lack of dependence upon the particular parameter adopted, of the integral I itself. F is a function of the two $N \times 1$ matrices X and X_τ and, under the change of parameter $\tau \to \tau'$, the second of these is multiplied by the positive factor $\tau_{\tau'}$ while the first is insensitive to this change of parameter. Thus, k being any positive number, the parametric integrand $F(X, X_\tau)$ is such that

$$F(X, kX_\tau) = kF(X, X_\tau)$$

We express this important quality of the parametric integrand F by the statement that F *is a positively homogeneous function, of degree* 1, *of the* $N \times 1$ *matrix* X_τ.

EXAMPLE

Denoting the rectangular Cartesian coordinates of a point in three-dimensional Euclidean space by (t, x^1, x^2), the formula that furnishes the arc-length I of any smooth curve in this space is

$$I = \int_a^b \{1 + (x_t^{\,1})^2 + (x_t^{\,2})^2\}^{\frac{1}{2}}\, dt = \int_a^b L\, dt$$

where $L = \{1 + (x_t^{\,1})^2 + (x_t^{\,2})^2\}^{\frac{1}{2}}$. Under the change of parameter $t \to \tau$ this appears in the form

$$I = \int_\alpha^\beta F\, d\tau$$

where $F = Lt_\tau = \{(t_\tau)^2 + (x_\tau^{\,1})^2 + (x_\tau^{\,2})^2\}^{\frac{1}{2}}$. Thus F is the magnitude, $(X_\tau^{\,*}X_\tau)^{\frac{1}{2}}$, of the 3×1 matrix

$$X_\tau = \begin{vmatrix} t \\ x^1 \\ x^2 \end{vmatrix}_\tau = \begin{vmatrix} t_\tau \\ x_\tau^{\,1} \\ x_\tau^{\,2} \end{vmatrix}$$

Observe that F, while a positively homogeneous function, of degree 1, of X_τ is not a homogeneous function, of degree 1, of X_τ; when X_τ is multiplied by a negative number k, F is multiplied by $-k = |k|$.

We shall assume from now on that the Lagrangian function L is a continuously differentiable function of the $N \times 1$ matrix X and of the $n \times 1$ matrix x_t. The derivative of L with respect to the $n \times 1$ velocity matrix x_t is a $1 \times n$ matrix which is known as the Lagrangian momentum matrix and which is denoted by p. Similarly the derivative of the parametric integrand F, which is a continuously differentiable function of the two $N \times 1$ matrices X and X_τ, with respect to the $N \times 1$ matrix X_τ, is a $1 \times N$ matrix P, which we term the parametric momentum matrix. In both of these differentiations the $N \times 1$ matrix X is supposed to be held fixed. The first element t_τ of X_τ appears in both the factors,

$$L = L(X, x_t) = L\left(X, \frac{x_\tau}{t_\tau}\right) \qquad \text{and} \qquad t_\tau$$

of F and so

$$P_1 = F_{t_\tau} = L_{t_\tau} t_\tau + L$$

Now

$$L_{t_\tau} = -L_{x_t}\frac{x_\tau}{t_\tau^2} = -\frac{p x_t}{t_\tau}$$

so that

$$P_1 = L - p x_t$$

The remaining elements of X_τ, namely, the elements of x_τ, appear only in the first factor L of F and so

$$L_{x_\tau^{j-1}} t_\tau, \qquad j = 2, \cdots, N,$$

$$= p_{j-1}\frac{1}{t_\tau} \cdot t_\tau = p_{j-1}$$

In words: The first element of the $1 \times N$ parametric matrix P is found by subtracting from the Lagrangian function L, the product of the $n \times 1$ velocity matrix x_t by the $1 \times n$ Lagrangian momentum matrix p; and the remaining $N - 1 = n$ elements of the parametric momentum matrix are those of the Lagrangian momentum matrix.

EXERCISE 1

Show that the elements of the parametric momentum matrix are positively homogeneous functions, of degree zero, of the $N \times 1$ matrix X_τ.

(*Hint*: Differentiate the relation $F(X, kX_\tau) = kF(X, X_\tau)$ with respect to X_τ.)

EXERCISE 2

Show that $PX_\tau = F$. (*Hint*: Differentiate the relation $F(X, kX_\tau) = kF(X, X_\tau)$ with respect to k and then set $k = 1$.)

EXERCISE 3

Show that, if $L = T - V$, where T is a homogeneous function of degree 2 of the $n \times 1$ velocity matrix x_t, and V does not involve x_t, then $P_1 = -(T + V)$. (*Hint*: $p = T_{x_t}$, $px_t = 2T$.)

2

Extremal Curves; The Euler-Lagrange Equation

The Lagrangian function L of a problem of Type 1 of the calculus of variations is a function of the $N \times 1$ matrix X and of the $n \times 1$ matrix x_t. We introduce the $(N + n) \times 1 = (2n + 1) \times 1$ matrix z whose first N elements are those of X and whose last n elements are those of x_t, and we regard the elements of z as the coordinates of a point in a space of $2n + 1$ dimensions. This space, which is known as the state space of the mechanical system whose Lagrangian function is L, is, like the time-coordinate space of the system, a numerical rather than a metrical space. We endow this numerical state-space with a Euclidean metric and we consider a region, i.e., an open, connected $(2n + 1)$ - dimensional domain D, in this $(2n + 1)$ - dimensional state-space over which $L = L(z)$ is, by hypothesis, a continuously differentiable function of z. If $X = X(t)$, $a \leqslant t \leqslant b$, is any smooth, or piecewise-smooth, curve C in the time-coordinate space, each point of C at which C is smooth furnishes a point of the state-space, the first N elements of the corresponding matrix z being those of $X(t)$ and the last n elements of z being those of x_t. The various points of the state-space which we obtain in this way constitute a curve Γ which we term the image of C in the state-space. We shall confine our attention to those piecewise-smooth curves C in the time-coordinate space whose images Γ in the state-space are covered by the region D over which $L = L(z)$ is, by hypothesis, a continuously differentiable function of z.

When our problem, of Type 1, of the calculus of variations is presented parametrically, our integral I appears as

$$I = \int_{\alpha}^{\beta} F(X, X_{\tau})\, d\tau$$

the curve C in the N-dimensional time-coordinate space along which I is evaluated being furnished by equations of the form

$$X = X(\tau), \qquad \alpha \leqslant \tau \leqslant \beta$$

The point z of our $(2n + 1)$-dimensional state-space, which is furnished by any point of C at which C is smooth, has as its first N coordinates the N elements of $X(\tau)$ and as its last n coordinates the n mutual ratios of the N elements of $X_\tau(\tau)$, it being assumed that $X_\tau(\tau)$ is not the zero $N \times 1$ matrix save, possibly, at a finite number of points of the interval $\alpha \leqslant \tau \leqslant \beta$. Thus z is independent of the particular parameter chosen to describe the curve C of integration. The parametric integrand $F(X, X_\tau)$ is a positively homogeneous function of degree 1 of the $N \times 1$ matrix X_τ, and the value of the integral I is independent of the parameter adopted to describe C.

In order to gain some idea as to how the integral I varies when the curve C along which it is evaluated is varied, we consider the following 1-parameter family of piecewise-smooth curves

$$C_s : X(\tau, s) = X(\tau) + sf(\tau); \qquad \alpha \leqslant \tau \leqslant \beta; \qquad -\delta \leqslant s \leqslant \delta$$

Here s is the parameter which names the various curves of the 1-parameter family, and we suppose that s varies over a closed interval which is centered at $s = 0$. We observe that C_s reduces to C when $s = 0$, and we express this fact by the statement that we have imbedded C in the 1-parameter family of curves C_s. $f(\tau)$ is any convenient $N \times 1$ matrix which is piecewise-smooth over $\alpha \leqslant \tau \leqslant \beta$ and which reduces to the zero $N \times 1$ matrix when $\tau = \alpha$ and when $\tau = \beta$, so that all the curves of our 1-parameter family have the same end-points. We suppose that the various images Γ_s, in the state-space, of the curves C_s of our 1-parameter family are all covered by the region D of our state-space over which L is, by hypothesis, a continuously differentiable function of z. Since

$$F_X = t_\tau L_X ; \qquad P_1 = L - px_t , \qquad P_j = p_{j-1} , \qquad j = 2, \cdots , N$$

it follows that F is, over D, a continuously differentiable function of the $2N \times 1$ matrix $Z = \begin{vmatrix} X \\ X_\tau \end{vmatrix}$ whose first N elements are those of X and whose last N elements are those of X_τ. When I is evaluated along any curve C_s of our 1-parameter family, its value is a function of s which is furnished by the formula

$$I(s) = \int_\alpha^\beta F(X(\tau) + sf(\tau), X_\tau(\tau) + sf_\tau(\tau))\, d\tau$$

$I(s)$ is, for each value of s in the interval $-\delta \leqslant s \leqslant \delta$, a differentiable function of s, its derivative being

$$I_s(s) = \int_\alpha^\beta F_X(X(\tau) + sf(\tau), X_\tau(\tau) + sf_\tau(\tau))f(\tau)\, d\tau$$
$$+ \int_\alpha^\beta P(X(\tau) + sf(\tau), X_\tau(\tau) + sf_\tau(\tau))f_\tau(\tau)\, d\tau$$

We denote by δI the differential of $I(s)$ at $s = 0$ and we term δI the variation of I. Thus δI is the product of $I_s(0)$ by ds, where ds is an arbitrary number (which we may take to be 1). Similarly, we denote by δX the differential, with respect to s, of $X(\tau, s)$ at $s = 0$, and by δX_τ the differential, with respect to s, of $X_\tau(\tau, s)$ at $s = 0$ so that

$$\delta X = ds \cdot f(\tau); \qquad \delta X_\tau = ds \cdot f_\tau(\tau)$$

We observe that

$$(\delta X)_\tau = \delta X_\tau$$

and we express this result by the statement that the order of variation and differentiation with respect to τ is, when these operations are applied to $X(\tau, s)$, immaterial. In this notation, then, we have

$$\delta I = \int_\alpha^\beta \{F_X(X(\tau), X_\tau(\tau))\delta X + P(X(\tau), X_\tau(\tau))\delta X_\tau\}\, d\tau$$

In order to simplify this expression we observe that $F_X(X(\tau), X_\tau(\tau))$ is continuous over $\alpha \leqslant \tau \leqslant \beta$ save, possibly, for a finite number of points, namely, the points which furnish the points of C at which x_τ is not defined. Thus the $1 \times N$ matrix function

$$G(\tau) = \int_\alpha^\tau F_X(X(\tau), X_\tau(\tau))\, d\tau$$

is defined and is continuous at all points of the interval $\alpha \leqslant \tau \leqslant \beta$. If $\tau = \tau_1$ is a point of the interval $\alpha \leqslant \tau \leqslant \beta$ at which x_t is not defined, both $G(\tau_1 - 0)$ and $G(\tau_1 + 0)$ exist and are equal, their common value being $G(\tau_1)$. Furthermore, at those points of C at which x_t is defined, $G_\tau = F_X$ so that $dG = d\tau \cdot G_\tau = d\tau \cdot F_X$. Thus the matrix product $F_X(X(\tau), X_\tau(\tau))\delta X$ may be integrated by parts to yield

$$\int_\alpha^\beta \{F_X(X(\tau), X_\tau(\tau))\delta X\}\, d\tau = (G\delta X)\,\Big|_\alpha^\beta - \int_\alpha^\beta \{G(\delta X)_\tau\}\, d\tau$$

and, since $\delta X = ds \cdot f(\tau)$ is zero, by hypothesis, at $\tau = \alpha$ and at $\tau = \beta$, this reduces to $-\int_\alpha^\beta \{G(\delta X)_\tau\} d\tau$. Hence, since $\delta X_\tau = (\delta X)_\tau$, δI appears as

$$\delta I = \int_\alpha^\beta \{(P - G)(\delta X)_\tau\}\, d\tau$$

It is clear that δI will be zero for all allowable choices of $f(\tau)$ if the $1 \times N$ matrix $P - G$ is constant along C; indeed, if this is the case, $\int_\alpha^\beta \{P - G)(\delta X)_\tau\} d\tau$ is of the form

$$\sum_{j=1}^{N} A_j \int_\alpha^\beta (\delta X^j)_\tau \, d^\tau = \sum_{j=1}^{N} A_j \, \delta X^j |_\alpha^\beta$$

where A_j, $j = 1, \cdots, N$, is constant, and this is zero, since $\delta X^j = f^j(\tau) ds$, $j = 1, \cdots, N$, is zero, by hypothesis, at $\tau = \alpha$ and at $\tau = \beta$. This sufficient condition for δI to be zero, for all allowable choices of $f(\tau)$, is also necessary. To see this we observe that δI may be written in the form $\int_\alpha^\beta \{(P - G - A)(\delta X)_\tau\} d\tau$, where A is any constant $1 \times N$ matrix, and we consider the $1 \times N$ matrix function, $H(\tau) = \int_\alpha^\tau (P - G - A) d\tau$. It is clear that $H(\alpha) = 0$, and we may determine A, by means of the formula $(\beta - \alpha)A = \int_\alpha^\beta (P - G) d\tau$, so that $H(\beta) = 0$. $H(\tau)$ is piecewise-smooth over $\alpha \leqslant \tau \leqslant \beta$, its derivative at any point τ of this interval, which furnishes a point of C at which C is smooth, being $P - G - A$. Hence we may take as our $N \times 1$ matrix $f(\tau)$ the transpose $H^*(\tau)$ of $H(\tau)$, and, when we do this, $(\delta X)_\tau = ds \cdot (P - G - A)^*$ at all the points of C at which C is smooth. Thus

$$\delta I = ds \int_\alpha^\beta \{(P - G - A)(P - G - A)^*\} \, d\tau$$

is the product by ds of the integral along C of the squared magnitude of the $1 \times N$ matrix $P - G - A$, and so, for δI to be zero, $P - G - A$ must be zero at all the points of the interval $\alpha \leqslant \tau \leqslant \beta$ at which it is continuous. Thus we have the following important result:

The necessary and sufficient condition that δI, when evaluated along C, be zero for all allowable choices of the $N \times 1$ matrix $f(\tau)$, is that $P = G + A$ at all the points of C at which C is smooth, the $1 \times N$ matrix A being constant along C.

Since $G(\tau)$ is differentiable, with derivative F_X, at all the points of the interval $\alpha \leqslant \tau \leqslant \beta$ which furnish smooth points of C it follows that at all smooth points of C,

$$P_\tau = F_X$$

This equation is known as the Euler-Lagrange equation for problems of Type 1 of the calculus of variations. We term any piecewise-smooth curve along which it holds an extremal curve of F or of the Lagrangian function L. The laws of mechanics for systems which possess a potential energy function may be stated as follows:

The paths, or curves in the time-coordinate space, traced by the mechanical system are extremal curves of the Lagrangian function L, or, equivalently, of the parametric integrand F.

In differentiating the relation $P = G + A$, in order to obtain the Euler-Lagrange equation $P_\tau = F_X$, we have lost the fact that A is constant along C, since the relation $P_\tau = F_X$ would hold if A were merely piecewise-constant along C. If τ_1 is a point of the interval $\alpha \leqslant \tau \leqslant \beta$ which furnishes a point of C at which x_t is not defined, we know that $G(\tau_1 - 0) = G(\tau_1 + 0)$ and this implies that $P(\tau_1 - 0) = P(\tau_1 + 0)$. The parametric momentum matrix $P(\tau)$ is not defined at τ_1 but, on assigning to it at τ_1 the common value of $P(\tau_1 - 0)$ and $P(\tau_1 + 0)$, we see that it is defined and continuous at all the points of the extremal curve C. This is a remarkable fact since the velocity matrix x_t is not defined at the points of C at which C fails to be smooth. The last n coordinates of P are those of the Lagrangian momentum matrix $p = L_{x_t}$ which we may regard as a function of x_t, the $N \times 1$ matrix X being held fixed. Let us denote by $L_{x_t^*}$ the $n \times 1$ matrix which is the transpose p^* of the Lagrangian momentum matrix p and let us suppose that $L_{x_t^*}$ is, over the region D of our $(2n + 1)$-dimensional state-space, a continuously differentiable function of the $n \times 1$ velocity matrix x_t. Then the Jacobian matrix $L_{x_t^* x_t}$ of $p^* = L_{x_t^*}$ with respect to x_t is a symmetric $n \times n$ matrix of which the element in the jth row and kth column is $L_{x_t^j x_t^k}$, $j, k = 1, \cdots, n$. If this matrix is nonsingular over D, the relation $p = L_{x_t}$ defines, over D, x_t as a function of X and p, and so the relation $p(\tau_1 - 0) = p(\tau_1 + 0)$, which is a consequence of the relation $P(\tau_1 - 0) = P(\tau_1 + 0)$, forces the equality $x_t(\tau_1 - 0) = x_t(\tau_1 + 0)$. In words:

Any extremal curve C of L is, when the n-dimensional matrix $L_{x_t^* x_t}$ exists and is continuous and nonsingular over D, not merely piecewise-smooth but actually smooth, x_t existing and being continuous at all the points of C.

If we assume, in addition, that the $n \times 1$ matrix $L_{x_t^*}$ is, over D, a continuously differentiable function of the $(2n + 1) \times 1$ matrix $z = \begin{vmatrix} X \\ x_t \end{vmatrix}$ and not merely of the $n \times 1$ matrix x_t, then the theory of implicit functions assures us that the function $x_t(X, p)$ of X and p which is defined implicitly by the formula $L_{x_t^*} = p^*$ is a continuously differentiable function of the $(2n + 1) \times 1$ matrix $\begin{vmatrix} X \\ p^* \end{vmatrix}$. Since this $(2n + 1) \times 1$ matrix is continuously differentiable along C, it follows that x_t is continuously differentiable along C so that x_{tt} exists and is continuous at all the points of C. In words:

Any extremal curve C of L is, when the $n \times (2n + 1)$ matrix $L_{x_t^* z}$ exists and is continuous over D (its n-dimensional submatrix $L_{x_t^* x_t}$ being nonsingular over D) not merely smooth but possessed of continuous curvature, x_{tt} being defined and continuous at all the points of C.

Along C the parametric momentum matrix P is continuously differentiable, its derivative being furnished by the Euler-Lagrange equation, $P_\tau = F_X$. The last n of the equations furnished by this $1 \times N$ matrix equation may be written as $p_t = L_x$ or, equivalently, on taking the transpose of this $1 \times n$ matrix equation, as

$$L_{x_t^* x_t} x_{tt} + L_{x_t^* X} X_t = L_{x^*}$$

This $n \times 1$ matrix equation, which furnishes x_{tt}, along any extremal curve of L, as a function of the $(2n + 1) \times 1$ matrix $z = \begin{vmatrix} X \\ x_t \end{vmatrix}$, is the Euler-Lagrange equation for extremals which possess continuous curvature. We shall from now on suppose that L possesses the following two properties which guarantee that all its extremals possess continuous curvature:

(1) $L_{x_t^* z}$ exists and is continuous over D.
(2) $L_{x_t^* x_t}$ is nonsingular over D.

Exercise 1

Show that if F or, equivalently, L does not involve any given one of the elements X^j say, of X then the corresponding element P_j of the parametric momentum matrix is constant along any extremal curve C of F.

Exercise 2

Show that the principle of conservation of energy holds for any mechanical system whose Lagrangian function does not involve t explicitly. (*Hint:* $P_1 = -(T + V)$.)

Note

In view of the result of this exercise, a mechanical system whose Lagrangian function does not involve t explicitly is termed conservative.

3

Lagrangian Functions Which Are Linear in x_t

When defining an extremal curve C of a given Lagrangian function L we imbedded C, which we assumed at the beginning to be merely piecewise-smooth, in a 1-parameter family of curves

$$X(\tau, s) = X(\tau) + sf(\tau), \qquad \alpha \leqslant \tau \leqslant \beta, -\delta \leqslant s \leqslant \delta$$

where $f(\tau)$ is any convenient $N \times 1$ matrix which is piecewise-smooth over $\alpha \leqslant \tau \leqslant \beta$ and which vanishes at $\tau = \alpha$ and at $\tau = \beta$. We found out later that if L satisfies some not very restrictive conditions, C must, if it is to qualify as an extremal curve of L, be smooth and, in addition, possess continuous curvature. Despite this fact, it is convenient to permit the comparison curves $X = X(\tau, s)$, $\alpha \leqslant \tau \leqslant \beta$, $-\delta \leqslant s \leqslant \delta$, to be only piecewise-smooth so that $f(\tau)$ may fail to be differentiable at a finite number of points of the interval $\alpha \leqslant \tau \leqslant \beta$. We do not impose additional restrictions upon our extremal curve if we imbed it in a family, $X = X(\tau, s)$, $\alpha \leqslant \tau \leqslant \beta$, $-\delta \leqslant s \leqslant \delta$, which does not involve the parameter s linearly, provided that the $N \times 1$ matrix function $X(\tau, s)$ is such that $\delta X_\tau = (\delta X)_\tau$ or, equivalently, that the two mixed second-order derivatives $X_{\tau s}$ and $X_{s\tau}$ exist and are equal save, possibly, for a finite number of points of the interval $\alpha \leqslant \tau \leqslant \beta$. This will certainly be the case if these derivatives exist and are continuous over the rectangle $\alpha \leqslant \tau \leqslant \beta$, $-\delta \leqslant s \leqslant \delta$, with the possible exception of a finite number of values of τ. We may also imbed our extremal curve C in a k - parameter family where k is any integer > 1. In this case s is a $k \times 1$ matrix, the parameter matrix of the family, and $X_{\tau s}$,