CHAPTER TWO

Calculus of Variations and Applications

2.1. Maxima and minima. Applications of the *calculus of variations* are concerned chiefly with the determination of maxima and minima of certain expressions involving unknown functions. Certain techniques involved are analogous to procedures in the differential calculus, which are briefly reviewed in this section.

An important problem in the differential calculus is that of determining maximum and minimum values of a function $y = f(x)$ for values of x in a certain interval (a, b). If in that interval $f(x)$ has a continuous derivative, it is recalled that a *necessary* condition for the existence of a maximum or minimum at a point x_0 *inside* (a, b) is that $dy/dx = 0$ at x_0. A *sufficient* condition that y be a maximum (or a minimum) at x_0, relative to values at neighboring points, is that, in addition, $d^2y/dx^2 < 0$ (or $d^2y/dx^2 > 0$) at that point.

If z is a function of two independent variables, say $z = f(x, y)$, in a region $\mathscr{R}$, and if the partial derivatives $\partial z/\partial x$ and $\partial z/\partial y$ exist and are continuous throughout $\mathscr{R}$, then *necessary* conditions that z possess a relative maximum or minimum at an interior point (x_0, y_0) of $\mathscr{R}$ are that $\partial z/\partial x = 0$ and $\partial z/\partial y = 0$ simultaneously at (x_0, y_0). These two requirements are equivalent to the single requirement that

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy = 0$$

at a point (x_0, y_0), for arbitrary values of both dx and dy. *Sufficient* conditions for either a maximum or a minimum involve certain inequalities among the second partial derivatives (see Problem 1).

More generally, a *necessary* condition that a continuously differentiable function $f(x_1, x_2, \cdots, x_n)$ of n variables $x_1, x_2, \cdots, x_n$ have a relative maximum or minimum value at an interior point of a region is that

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n = 0 \tag{1}$$

at that point, for all *permissible* values of the differentials $dx_1, \cdots, dx_n$.

At a point satisfying (1) the function f is said to be *stationary*.

If the n variables are all independent, the n differentials can be assigned arbitrarily, and it follows that (1) then is equivalent to the n conditions

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \cdots = \frac{\partial f}{\partial x_n} = 0. \tag{2}$$

Sufficient conditions that values of the variables satisfying (1) or (2) actually determine maxima (or minima) involve certain inequalities among the higher partial derivatives (see Problem 1).

Suppose, however, that the n variables are *not* independent, but are related by, say, N conditions each of the form

$$\phi_k(x_1, \cdots, x_n) = 0.$$

Then, at least theoretically, these N equations generally can be solved to express N of the variables in terms of the $n - N$ remaining variables, and hence to express f and df in terms of $n - N$ *independent* variables and their differentials. Alternatively, N linear relations among the n *differentials* can be obtained by differentiation. These conditions permit the expression of N of the *differentials* as linear combinations of the differentials of the $n - N$ independent variables. If (1) is expressed in terms of these differentials, their coefficients must then vanish, giving $n - N$ conditions for stationary values of f which supplement the N constraint conditions.

A procedure which is often still more convenient in this case consists of the introduction of the so-called *Lagrange multipliers*. To illustrate their use, we consider here the problem of obtaining stationary values of $f(x, y, z)$,

$$df \equiv f_x\, dx + f_y\, dy + f_z\, dz = 0, \tag{3}$$

subject to the two constraints

$$\phi_1(x, y, z) = 0, \tag{4a}$$

$$\phi_2(x, y, z) = 0. \tag{4b}$$

Since the three variables x, y, z must satisfy the two auxiliary conditions (4a,b), only one variable can be considered as independent. Equations (4a,b) imply the differential relations

$$\phi_{1x}\, dx + \phi_{1y}\, dy + \phi_{1z}\, dz = 0, \tag{5a}$$

$$\phi_{2x}\, dx + \phi_{2y}\, dy + \phi_{2z}\, dz = 0. \tag{5b}$$

The procedure outlined above would consist of first solving (5a,b) for, say, dx and dy in terms of dz (if this is possible) and of introducing the results into (3), to give a result of the form

$$df = (\cdots)\, dz = 0.$$

Since dz can be assigned arbitrarily, the vanishing of the indicated expression in parentheses in this form is the desired condition that f be stationary when (4a,b) are satisfied.

As an alternative procedure, we first multiply (5a) and (5b) respectively by the quantities λ_1 and λ_2, to be specified presently, and add the results to (3). Since the right-hand members are all zeros, there follows

$$(f_x + \lambda_1\phi_{1x} + \lambda_2\phi_{2x})\, dx + (f_y + \lambda_1\phi_{1y} + \lambda_2\phi_{2y})\, dy + (f_z + \lambda_1\phi_{1z} + \lambda_2\phi_{2z})\, dz = 0, \qquad (6)$$

for arbitrary values of λ_1 *and* λ_2. Now let λ_1 and λ_2 be determined so that two of the parentheses in (6) vanish.* Then the differential multiplying the remaining parenthesis can be arbitrarily assigned, and hence that parenthesis must also vanish. Thus we must have

$$\left.\begin{aligned} \frac{\partial f}{\partial x} + \lambda_1 \frac{\partial \phi_1}{\partial x} + \lambda_2 \frac{\partial \phi_2}{\partial x} &= 0, \\ \frac{\partial f}{\partial y} + \lambda_1 \frac{\partial \phi_1}{\partial y} + \lambda_2 \frac{\partial \phi_2}{\partial y} &= 0, \\ \frac{\partial f}{\partial z} + \lambda_1 \frac{\partial \phi_1}{\partial z} + \lambda_2 \frac{\partial \phi_2}{\partial z} &= 0 \end{aligned}\right\}. \qquad (7a,b,c)$$

Equations (7a,b,c) and (4a,b) comprise five equations determining x, y, z and λ_1, λ_2. The quantities λ_1 and λ_2 are known as *Lagrange multipliers*. Their introduction frequently simplifies the relevant algebra in problems of the type just considered. In many applications they are found to have physical significance as well. We notice that *the conditions* (7) *are the conditions that* $f + \lambda_1\phi_1 + \lambda_2\phi_2$ *be stationary when no constraints are present*.

The procedure outlined is applicable without modification to the general case of n variables and $N < n$ constraints.

In illustration of the method, we attempt to determine the point on the curve of intersection of the surfaces

$$z = xy + 5, \qquad x + y + z = 1 \qquad (8a,b)$$

which is nearest the origin. Thus, we must minimize the quantity

$$f = x^2 + y^2 + z^2$$

* It can be shown that if this were not possible, then the functions ϕ_1 and ϕ_2 would be *functionally dependent*, so that the two constraints (4a) and (4b) would be either equivalent or incompatible.

subject to the two constraints (8a,b). With

$$\phi_1 = z - xy - 5, \qquad \phi_2 = x + y + z - 1,$$

equations (7a,b,c) take the form

$$\left.\begin{aligned} 2x - \lambda_1 y + \lambda_2 &= 0, \\ 2y - \lambda_1 x + \lambda_2 &= 0, \\ 2z + \lambda_1 \quad + \lambda_2 &= 0 \end{aligned}\right\}. \tag{9a,b,c}$$

The elimination of λ_1 and λ_2 from equations (9a,b,c) yields the two alternatives

$$x + y - z + 1 = 0 \qquad or \qquad x = y. \tag{10a,b}$$

The simultaneous solution of (8a,b) and (10a) leads to the coordinates of the two points (2, −2, 1) and (−2, 2, 1), which are each three units distant from the origin, whereas the equations (8a,b) and (10b) have *no* real common solution. Geometrical considerations indicate that there is indeed at least one point nearest the origin; since the two points obtained are *necessarily* the only possible ones, they must accordingly be the points required.

As an illustration closely related to certain topics in Chapter 1, we may seek those points on a central quadric surface

$$\phi \equiv a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{23}yz + 2a_{13}xz = \text{constant}$$

for which distance from the origin is maximum or minimum relative to neighboring points. We are thus to render the form

$$f \equiv x^2 + y^2 + z^2$$

stationary, subject to the constraint $\phi = $ constant. Here, if we denote the Lagrange multiplier by $-1/\lambda$, the requirement that $\phi - \lambda f$ be stationary leads to the conditions

$$\left.\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= \lambda x, \\ a_{12}x + a_{22}y + a_{23}z &= \lambda y, \\ a_{13}x + a_{23}y + a_{33}z &= \lambda z \end{aligned}\right\}.$$

This set of equations comprises a characteristic-value problem of the type discussed in Section 1.12. Each "characteristic value" of λ, for which a nontrivial solution exists, leads to the three coordinates of one or more *points* (x, y, z), determined within a common arbitrary multiplicative factor which is available for the satisfaction of the equation of the surface. Section 1.21 shows that it is always possible to rotate the coordinate axes in such a way that each new axis coincides with the direction from the origin to such a point, and that the equation of the surface, referred to the new axes, then involves only squares of the new coordinates. That is, the new axes (which coincide with the "characteristic vectors" of the problem) are the *principal axes* of the quadric surface. The characteristic values of λ are inversely proportional to the squares of the semiaxes. Repeated roots of the characteristic equation correspond to surfaces of revolution, in which cases the new axes can be so chosen in infinitely many ways, while zero roots correspond to surfaces which extend infinitely far from the origin.

The basic problem in the *calculus of variations* is to determine a *function* such that a certain definite integral involving that function and certain of its derivatives takes on a maximum or minimum value. The elementary part of the theory is concerned with a *necessary* condition (generally in the form of a differential equation with boundary conditions) which the required function must satisfy. To show mathematically that the function obtained actually maximizes (or minimizes) the integral is much more difficult than in the corresponding problems of differential calculus. *Sufficient* conditions are developed in more advanced works. In physically motivated problems, such additional considerations frequently may be avoided.

As an example of a problem of this sort, we notice that in order to determine the surface of revolution, obtained by rotating about the x axis a curve passing through two given points (x_1, y_1) and (x_2, y_2), which has minimum surface area, we must determine the function $y(x)$ which specifies the curve to be revolved, in such a way that the integral

$$I = 2\pi \int_{x_1}^{x_2} y(1 + y'^2)^{1/2}\, dx$$

is a minimum, and also so that $y(x_1) = y_1$ and $y(x_2) = y_2$. Here it is assumed that y_1 and y_2 are nonnegative.

In most cases it is to be required that the function and the derivatives explicitly involved be continuous in the region of definition.

2.2. The simplest case. We now consider the problem of determining a continuously differentiable function $y(x)$ for which the integral

$$I = \int_{x_1}^{x_2} F(x, y, y')\, dx \tag{11}$$

takes on a maximum or minimum value,* and which satisfies the prescribed end conditions

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

To fix ideas, we may suppose that I is to be *minimized*.

Suppose that $y(x)$ is the actual minimizing function, and choose *any* continuously differentiable function $\eta(x)$ which *vanishes* at the end points

* We suppose that F has continuous second partial derivatives with respect to its three arguments

$x = x_1$ and $x = x_2$. Then for any constant ϵ the function $y(x) + \epsilon\eta(x)$ will satisfy the end conditions (Figure 2.1). The integral

$$I(\epsilon) = \int_{x_1}^{x_2} F(x, y + \epsilon\eta, y' + \epsilon\eta')\, dx, \tag{12}$$

obtained by replacing y by $y + \epsilon\eta$ in (11), is then a function of ϵ, once y and η are assigned, which takes on its minimum value when $\epsilon = 0$. But this is possible only if

$$\frac{dI(\epsilon)}{d\epsilon} = 0 \quad when \quad \epsilon = 0. \tag{13}$$

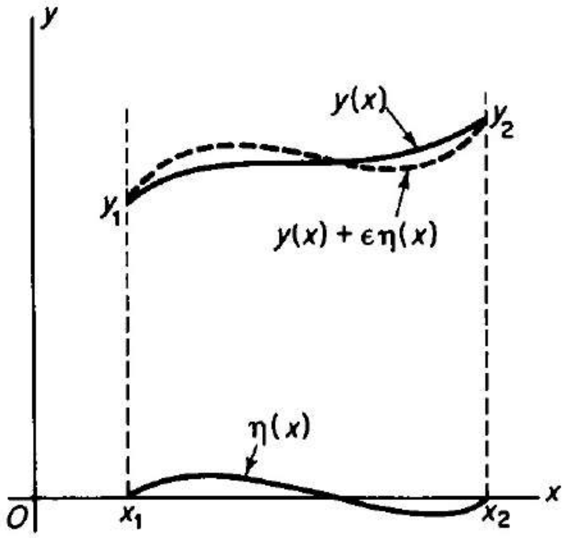

FIGURE 2.1

If we denote the integrand in (12) by $\tilde{F}$,

$$\tilde{F} = F(x, y + \epsilon\eta, y' + \epsilon\eta'),$$

and notice that

$$\frac{d\tilde{F}}{d\epsilon} = \frac{\partial \tilde{F}}{\partial y}\eta + \frac{\partial \tilde{F}}{\partial y'}\eta',$$

we obtain from (12) the result

$$\frac{dI(\epsilon)}{d\epsilon} = \int_{x_1}^{x_2} \left(\frac{\partial \tilde{F}}{\partial y}\eta + \frac{\partial \tilde{F}}{\partial y'}\frac{d\eta}{dx}\right) dx,$$

by differentiating under the integral sign. Finally, since $\tilde{F} \to F$ when $\epsilon \to 0$, and the same is true of the partial derivatives, the necessary condition (13) takes the form

$$\int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\frac{d\eta}{dx}\right) dx = 0. \tag{14}$$

The next step in the development consists of integrating the second term by parts, to transform (14) to the condition

$$\int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y}\eta - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\eta\right] dx + \left[\frac{\partial F}{\partial y'}\eta(x)\right]_{x_1}^{x_2} = 0. \tag{15}$$

But since $\eta(x)$ vanishes at the end points, by assumption, the integrated terms vanish and (15) becomes

$$\int_{x_1}^{x_2} \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\eta\, dx = 0. \tag{16}$$

Finally, since $\eta(x)$ is arbitrary, we conclude that its coefficient in (16) must vanish identically over (x_1, x_2). For if this were not so we could choose a continuously differentiable function $\eta(x)$, which vanishes at the ends of the interval, in such a way that the (continuous) integrand in (16) is positive whenever it is not zero,* and a contradiction would be obtained.

The end result is that if $y(x)$ minimizes (or maximizes) the integral (11), it must satisfy the *Euler equation*

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) - \frac{\partial F}{\partial y} = 0. \tag{17a}$$

Here the partial derivatives $\partial F/\partial y$ and $\partial F/\partial y'$ have been formed by treating x, y, and y' as independent variables. Since $\partial F/\partial y'$ is, in general, a function of x explicitly and also implicitly through y and $y' = dy/dx$, the first term in (17a) can be written in the expanded form

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial y'}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial y'}\right)\frac{dy}{dx} + \frac{\partial}{\partial y'}\left(\frac{\partial F}{\partial y'}\right)\frac{dy'}{dx}.$$

Thus (17a) is equivalent to the equation

$$F_{y'y'}\frac{d^2y}{dx^2} + F_{y'y}\frac{dy}{dx} + (F_{y'x} - F_y) = 0. \tag{17b}$$

This equation is of second order in y unless $F_{y'y'} = \partial^2 F/\partial y'^2 \equiv 0$, so that in general two constants are available for the satisfaction of the end conditions.

It is useful to notice that (17b) is equivalent to the form

$$\frac{1}{y'}\left[\frac{d}{dx}\left(F - \frac{\partial F}{\partial y'}\frac{dy}{dx}\right) - \frac{\partial F}{\partial x}\right] = 0, \tag{17c}$$

as can be verified by expansion (see also Problem 7). From this result it follows that if F does not involve x *explicitly* a first integral of Euler's equation is

$$F - y'\frac{\partial F}{\partial y'} = C \quad \textit{if} \quad \frac{\partial F}{\partial x} \equiv 0, \tag{18a}$$

while (17a) shows that if F does not involve y explicitly a first integral is

$$\frac{\partial F}{\partial y'} = C \quad \textit{if} \quad \frac{\partial F}{\partial y} \equiv 0. \tag{18b}$$

Solutions of Euler's equation are known as *extremals* of the problem considered. In general, they comprise a two-parameter family of functions in the case just treated.

An extremal which satisfies the appropriate end conditions is often called a *stationary function* of the variational problem, and is said to make the relevant integral *stationary*, whether or not it also makes the integral *maximum* or *minimum*, relative to all slightly varied admissible functions.

* This fact, which is intuitively plausible, can be proved analytically.

Thus, by definition, the integral I of equation (11) will be said to be *stationary* when $y(x)$ is so determined that equation (15) holds for every permissible function $\eta(x)$.*

2.3. Illustrative examples. In Section 2.1 it was pointed out that to find the minimal surface of revolution passing through two given points it is necessary to minimize the integral

$$\frac{I}{2\pi} = \int_{x_1}^{x_2} y(1 + y'^2)^{1/2}\,dx. \tag{19}$$

With $F = y(1 + y'^2)^{1/2}$, the Euler equation (17a) becomes

$$\frac{d}{dx}\left[\frac{yy'}{(1 + y'^2)^{1/2}}\right] - (1 + y'^2)^{1/2} = 0$$

or, after a reduction or use of (17b),

$$yy'' - y'^2 - 1 = 0. \tag{20}$$

Following the usual procedure for solving equations of this type, we set

$$y' = p, \qquad y'' = \frac{dp}{dx} = p\,\frac{dp}{dy},$$

so that (20) becomes

$$py\,\frac{dp}{dy} = p^2 + 1.$$

This equation is separable, and is integrated to give

$$y = c_1(1 + p^2)^{1/2} = c_1\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2},$$

as would be obtained more directly by use of (18a), since here F does not explicitly involve x. There follows

$$\frac{dy}{dx} = \left(\frac{y^2}{c_1^{\,2}} - 1\right)^{1/2},$$

and hence finally

$$y = c_1 \cosh\left(\frac{x}{c_1} + c_2\right). \tag{21}$$

Thus, as is well known, the required minimal surface (if it exists) must be obtained by revolving a catenary. It then remains to be seen whether the arbitrary constants c_1 and c_2 can indeed be so chosen that the curve (21) passes through any two assigned points in the upper half plane.

* The usage of the terms "extremal" and "stationary function" varies within the literature.

The determination of these constants is found to involve the solution of a transcendental equation which possesses two, one, or no solutions, depending upon the prescribed values $y(x_1)$ and $y(x_2)$. In particular, it is found that all curves representing (21) and passing through *one* prescribed point $P_1(x_1, y_1)$ are tangent to a certain curve $\mathscr{C}_1$ which passes through the point $(x_1, 0)$, and that no such curve *crosses* $\mathscr{C}_1$. When the second prescribed point $P_2(x_2, y_2)$ is separated from P_1 by $\mathscr{C}_1$, there is accordingly *no* curve representing (21) passing through *both* points, and hence *no* admissible minimal surface of revolution. When P_1 and P_2 are on the same side of $\mathscr{C}_1$, there are *two* "stationary curves," the shorter of which generates a minimal surface. Finally, when P_2 is *on* $\mathscr{C}_1$, there is *one* stationary curve but the surface of revolution which it generates is *not* minimal.

The situations in which there is no admissible minimizing curve are those in which smaller and smaller areas are generated by rotating curves which more and more nearly approach a broken line consisting of segments from (x_1, y_1) to $(x_1, 0)$ to $(x_2, 0)$ to (x_2, y_2), and in which that *unattainable* limiting area is smaller than the area generated by any admissible curve.

Physically, the problem can be interpreted as that of determining the shape of a *soap film* connecting parallel circular wire hoops of radii y_1 and y_2, perpendicular to the x axis, with centers at $(x_1, 0)$ and $(x_2, 0)$. The exceptional cases arise when the separation $x_2 - x_1$ is increased to or beyond the point where the film no longer can *join* the hoops, but breaks into two parts, each then spanning a hoop in its plane.

The classical "elementary" application of the calculus of variations consists of *proving* mathematically that the shortest distance between two points in a plane is a straight line. If the points, in the xy plane, are (x_1, y_1) and (x_2, y_2) and if the equation of the minimizing curve is $y = y(x)$, we are then to minimize

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

Since here $F = (1 + y'^2)^{1/2}$ does not involve either x or y explicitly, either of the forms (18a, b) can be used to give a first integral of Euler's equation directly. However, here it is easier to use the form (17b) to deduce that $y'' = 0$ and consequently $y = c_1 x + c_2$. From this result we can conclude that *if a minimizing curve exists and if it can be specified by an equation of the form $y = y(x)$, then that curve necessarily must be a straight line.* It is clear that the case in which $x_1 = x_2$ is exceptional, and must be treated separately.

In the preceding examples no proof was given that the stationary function obtained actually possesses the required minimizing property. Such considerations comprise most of the less elementary theory of the calculus of variations. In a great number of physically motivated problems it is intuitively clear that a minimizing function does indeed *exist*. Then if the present methods and their extensions show that only the particular function obtained could possibly be the minimizing function, the problem can be considered as solved for practical purposes. If several alternatives (stationary functions) are determined, direct calculation will show which one actually leads to the smaller value of the quantity to be minimized.

In many practical situations the stationary functions are of importance whether or not they maximize or minimize the relevant integral. This fact is illustrated, for example, in Section 2.9.

2.4. Natural boundary conditions and transition conditions. When the value of the unknown function $y(x)$ is *not* preassigned at one or both of the end points $x = x_1, x_2$ the difference $\epsilon\eta(x)$ between the true function $y(x)$ and the varied function $y(x) + \epsilon\eta(x)$ need not vanish there. However, the left-hand member of (15) must vanish when $y(x)$ is identified with the minimizing (or maximizing) function, for *all* permissible variations $\epsilon\eta(x)$. Thus it must vanish, *in particular*, for all variations which *are* zero at both ends. For all such η's the second term in (15) is zero and equation (16) again follows, and yields the Euler equation (17) as before.

Hence the first term in (15) must be zero for *all* permissible η's, since the *coefficient* of η in the integrand must be zero. Thus it follows that the second term in (15) must itself vanish,

$$\left[\frac{\partial F}{\partial y'}\,\eta(x)\right]_{x=x_2} - \left[\frac{\partial F}{\partial y'}\,\eta(x)\right]_{x=x_1} = 0, \tag{22}$$

for all permissible values of $\eta(x_2)$ and $\eta(x_1)$. If $y(x)$ is not preassigned at either end point, then $\eta(x_1)$ and $\eta(x_2)$ are both completely arbitrary and we conclude that their coefficients in (22) must each vanish, yielding the conditions

$$\left[\frac{\partial F}{\partial y'}\right]_{x=x_1} = 0, \qquad \left[\frac{\partial F}{\partial y'}\right]_{x=x_2} = 0 \tag{23a,b}$$

which must be satisfied instead.

The requirements that (23a) hold when $y(x_1)$ is not given, and that (23b) hold when $y(x_2)$ is not given, are often called the *natural boundary conditions* of the problem.* If, for example, $y(x_1)$ were preassigned as y_1 whereas $y(x_2)$ were not given in advance, then the relevant end conditions would be $y(x_1) = y_1$ and $(\partial F/\partial y')_{x=x_2} = 0$, and (23a) would *not* apply.

In some situations the integrand F in (11) is such that one or both of the terms $\partial F/\partial y$ and $d(\partial F/\partial y')/dx$ are discontinuous at one or more points inside the interval (x_1, x_2), but the conditions assumed in the preceding section are satisfied in the subintervals separated by these points. To illustrate the treatment of such cases, we suppose here that there is only *one* point of

* In some references, only the conditions (23a,b) themselves are called the "natural boundary conditions."

discontinuity, at $x = c$. Then the integral (11) must be expressed as the sum of integrals over $(x_1, c-)$ and $(c+, x_2)$ before the steps leading to (15) are taken, and equation (15) is replaced by the relation

$$\int_{x_1}^{c-}\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\eta\,dx + \int_{c+}^{x_2}\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\eta\,dx + \left[\frac{\partial F}{\partial y'}\,\eta(x)\right]_{x_1}^{c-} + \left[\frac{\partial F}{\partial y'}\,\eta(x)\right]_{c+}^{x_2} = 0. \qquad (24)$$

If we require that the minimizing (or maximizing) function $y(x)$ be *continuous* at $x = c$, and accordingly require that all admissible functions $y(x) + \epsilon\eta(x)$ have the same property, it follows that

$$\eta(c+) = \eta(c-) = \eta(c),$$

so that (24) can be written in the form

$$\int_{x_1}^{c-}\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\eta\,dx + \int_{c+}^{x_2}\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\eta\,dx + \left[\frac{\partial F}{\partial y'}\right]_{x_2}\eta(x_2) - \left[\frac{\partial F}{\partial y'}\right]_{x_1}\eta(x_1) - \left\{\left[\frac{\partial F}{\partial y'}\right]_{c+} - \left[\frac{\partial F}{\partial y'}\right]_{c-}\right\}\eta(c) = 0. \qquad (25)$$

Hence we may deduce that the Euler equation (17) must hold in each of the subintervals (x_1, c) and (c, x_2), that $\partial F/\partial y'$ must vanish at any end point $x = x_1$ or $x = x_2$ where y is not prescribed, as before, and also that the *natural transition conditions*

$$y(c+) = y(c-), \qquad \lim_{x\to c+}\frac{\partial F}{\partial y'} = \lim_{x\to c-}\frac{\partial F}{\partial y'} \qquad (26a,b)$$

must be satisfied at the point $x = c$. Whereas (26a) represents the requirement that y itself be continuous at $x = c$, the condition (26b) may demand that the *derivative* y' be *discontinuous* at that point.

To illustrate the preceding considerations, we consider the determination of stationary functions associated with the integral

$$I = \int_0^1 (Ty'^2 - \rho\omega^2 y^2)\,dx, \qquad (27)$$

where T, ρ, and ω are given constants or functions of x. The Euler equation is

$$\frac{d}{dx}\left(T\frac{dy}{dx}\right) + \rho\omega^2 y = 0, \qquad (28)$$

regardless of what (if anything) is prescribed in advance at the ends of the interval.

Thus, in particular, when T, ρ, and ω are *positive constants*, the extremals are of the form

$$y = c_1 \cos \alpha x + c_2 \sin \alpha x \qquad \left(\alpha^2 = \frac{\rho\omega^2}{T}\right), \tag{29}$$

where c_1 and c_2 are constants. When the conditions

$$y(0) = 0, \qquad y(1) = 1$$

are prescribed, there follows $c_1 = 0$, $c_2 \sin \alpha = 1$ and hence

$$y = \frac{\sin \alpha x}{\sin \alpha} \qquad (\alpha \neq \pi, 2\pi, \cdots). \tag{30}$$

When the condition

$$y(0) = 1$$

is prescribed, but $y(1)$ is *not* preassigned, the appropriate condition at $x = 1$ follows from (23b), with $F = Ty'^2 - \rho\omega^2 y^2$, in the form

$$Ty'(0) = 0$$

and hence

$$y = \frac{\cos \alpha x}{\cos \alpha} \qquad \left(\alpha \neq \frac{\pi}{2}, \frac{3\pi}{2}, \cdots\right). \tag{31}$$

When neither $y(0)$ nor $y(1)$ is prescribed, the conditions (23a,b) require

$$Ty'(0) = Ty'(1) = 0$$

and hence

$$y = \begin{cases} 0 & (\alpha \neq \pi, 2\pi, \cdots), \\ c_1 \cos \alpha x & (\alpha = \pi, 2\pi, \cdots), \end{cases} \tag{32}$$

where c_1 is arbitrary. In the exceptional cases noted in (30) and (31), no stationary function exists. The limiting cases in which $\alpha = 0$ must be treated separately, since (29) is incomplete in that case.

If $T = T_1$ and $\rho = \rho_1$ when $0 \leqq x < c$ whereas $T = T_2$ and $\rho = \rho_2$ when $c < x \leqq 1$, where T_1, T_2, ρ_1, ρ_2, and ω are positive constants, and if the conditions

$$y(0) = 0, \qquad y(1) = 1$$

are prescribed, there follows

$$y = \begin{cases} c_1 \cos \alpha_1 x + c_2 \sin \alpha_1 x & (0 \leqq x < c), \\ d_1 \cos \alpha_2 x + d_2 \sin \alpha_2 x & (c < x \leqq 1), \end{cases}$$

where $\alpha_i^2 = \rho_i \omega^2 / T_i$. The natural transition conditions (26a,b) also give

$$\lim_{x \to c-} y(x) = \lim_{x \to c+} y(x), \qquad T_1 \lim_{x \to c-} y'(x) = T_2 \lim_{x \to c+} y'(x).$$

Thus we have four conditions which are to be satisfied by the four constants of integration, and a stationary function is determined provided that α does not take on one of a certain infinite set of exceptional values (which correspond to the vanishing of the determinant of a certain coefficient matrix).

2.5. The variational notation. We next introduce the notation of "variations" in order to establish more clearly the analogy between the calculus of variations and the differential calculus.

Suppose that we consider a set $\mathscr{S}$ of functions satisfying certain conditions. For example, we might define $\mathscr{S}$ to be the set of all functions of a single variable x which possess a continuous first derivative at all points in an interval $a \leq x \leq b$. Then any quantity which takes on a specific numerical value corresponding to each function in $\mathscr{S}$ is said to be a *functional* on the set $\mathscr{S}$.

In illustration, we may speak of the quantities

$$I_1 = \int_a^b y(x)\,dx, \qquad I_2 = \int_a^b \{y(x)y''(x) - [y'(x)]^2\}\,dx$$

as functionals, since corresponding to any function $y(x)$ for which the indicated operations are defined each quantity has a definite numerical value.

With the above definition, it is proper also to speak of such quantities as $f[y(x)]$ and $g[x, y(x), y'(x), \cdots, y^{(n)}(x)]$ as functionals in those cases when the variable x is considered as fixed in a given discussion and the function $y(x)$ is varied.

In Section 2.2, we considered an integrand of the form

$$F = F(x, y, y')$$

which *for a fixed value of* x depends upon the *function* $y(x)$ and its derivative. We then changed the function $y(x)$, to be determined, into a new function $y(x) + \epsilon\eta(x)$. The change $\epsilon\eta(x)$ in $y(x)$ is called the *variation* of y and is conventionally denoted by δy,

$$\delta y = \epsilon\eta(x). \tag{33}$$

Corresponding to this change in $y(x)$, for a fixed value of x, the functional F changes by an amount ΔF, where

$$\Delta F = F(x, y + \epsilon\eta, y' + \epsilon\eta') - F(x, y, y'). \tag{34}$$

If the right-hand member is expanded in powers of ϵ, there follows

$$\Delta F = \frac{\partial F}{\partial y}\epsilon\eta + \frac{\partial F}{\partial y'}\epsilon\eta' + (\text{terms involving higher powers of } \epsilon). \tag{35}$$

In analogy with the definition of the *differential*, the first two terms in the right-hand member of (35) are *defined* to be the *variation of F*,

$$\delta F = \frac{\partial F}{\partial y}\epsilon\eta + \frac{\partial F}{\partial y'}\epsilon\eta'. \tag{36}$$