Books on diffusion, heat flow, and transport theory:

Chandrasekhar, S.: "Radiative Transfer," Oxford, New York, 1950.
Chapman, S., and T. G. Cowling: "Mathematical Theory of Non-uniform Gases," Cambridge, New York, 1939.
Fowler, R. H.: "Statistical Mechanics," Cambridge, New York, 1936.
Hopf, E.: "Mathematical Problems of Radiative Equilibrium," Cambridge, New York, 1934.
Lorentz, H. A.: "Theory of Electrons," B. G. Teubner, Leipzig, 1909.

Texts on electromagnetic theory, particularly on fundamental concepts:

Abraham, M., and R. Becker: "Classical Theory of Electricity and Magnetism," Blackie, Glasgow, 1932.
Stratton, J. A.: "Electromagnetic Theory," McGraw-Hill, New York, 1941.
Van Vleck, J. H.: "Theory of Electric and Magnetic Susceptibilities," Chaps. 1–4, Oxford, New York, 1932.

Discussions of the fundamental principles of quantum mechanics from various points of view:

Bohm, D.: "Quantum Theory," Prentice-Hall, New York, 1951.
De Broglie, L.: "L'Electron magnetique," Hermann & Cie, Paris, 1945.
Dirac, P. A. M.: "Principles of Quantum Mechanics," Oxford, New York, 1935.
Jordan, P.: "Anschauliche Quantentheorie," Springer, Berlin, 1936.
Kemble, E. C.: "Fundamental Principles of Quantum Mechanics," McGraw-Hill, New York, 1937.
Kramers, H. A.: "Grundlagen der Quantentheorie," Akademische Verlagsgesellschaft m.b.H., Leipzig, 1938.
Sommerfeld, A.: "Wellenmechanik," Vieweg, Brunswick, 1939, reprinted, Ungar, New York, 1946.
Van der Waerden, B. L.: "Gruppentheoretische Methode in der Quantenmechanik," Springer, Berlin, 1932.
Von Neumann, J.: "Mathematische Grundlagen der Quantenmechanik," Springer, Berlin, 1932.

CHAPTER 3

Fields and the Variational Principle

The use of superlatives enables one to express in concise form a general principle covering a wide variety of phenomena. The statements, for instance, that a straight line is the shortest distance between two points or that a circle is the shortest line which encloses a given area, are deceptively simple ways of defining geometrical entities. To say that electric current distributes itself in a network of resistors so that the least power is converted into heat is a description of direct-current flow which encompasses many individual cases without the use of mathematical complexities (though the complexities inevitably intrude when the general principle is applied to the individual case). The statement that a physical system so acts that some function of its behavior is least (or greatest) is often both the starting point for theoretical investigation and the ultimate distillation of all the relationships between facts in a large segment of physics.

The mathematical formulation of the superlative is usually that the integral of some function, typical of the system, has a smaller (or else larger) value for the actual performance of the system than it would have for any other imagined performance subject to the same very general requirements which serve to particularize the system under study. We can call the integrand L; it is a function of a number of independent variables of the system (coordinates, field amplitudes, or other quantities) and of the derivatives of these variables with respect to the parameters of integration (velocities or gradients of fields, etc.). If the variables are $\varphi_1, \ldots, \varphi_n$, the parameters $x_1, \ldots, x_m$ and the derivatives $\partial\varphi_r/\partial x_s = \varphi_{rs}$, then the integral which is to be minimized is

$$\mathfrak{L} = \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} L\left(\varphi, \frac{\partial\varphi}{\partial x}, x\right) dx_1 \cdots dx_m \tag{3.1.1}$$

From the minimization of this function we can obtain the partial differential equations governing the φ's as functions of the x's and many other things. This process of obtaining the φ's is called the *variational method.*

In the present chapter we shall first indicate in a general way how

the variational method can be used to obtain equations for the variables involved, then discuss in more detail the variational principles of classical dynamics, because they provide a well-explored example of this technique and its utility; and then we shall proceed to apply the method to the various types of fields which are to be studied in this book.

3.1 *The Variational Integral and the Euler Equations*

The integrand L of the integral to be minimized (or maximized) will be called the *Lagrange density* of the system. It is a *function of functions* of the basic parameters of the system. In the case of classical dynamics, for instance, the parameter is the time and the functions are the coordinates and velocities, at various times, of the configuration of the system as it moves in conformity with the applied forces and the initial conditions; in the case of fields the basic parameters are the coordinates locating every point where the field is to be measured, and the functions are the various components of the field and their gradients, which are determined, as functions of these coordinates, by the distribution of the various "sources" (or charges) in space and by the boundary conditions.

Thus when we require that the integral of L should be minimized (or maximized), we mean that the functions in terms of which L is expressed (the coordinates and velocities or fields and gradients) are to be adjusted, at every value of the parameters, so that the integral of L has the least possible value. We wish so to adjust the functions φ that the integral of L, a function of the φ's and $\partial\varphi/\partial x$'s, is as small as it can become, subject to the conditions imposed by the state of the system.

In order to solve the problem we must first make the step from a variational requirement on an integral of L to a set of equations which determine the best values of the functions φ.

The Euler Equations. Before we can make even this step we must be more specific about what we mean by "minimizing the integral" and "best values of the functions." To do this, suppose we arbitrarily choose a functional form for each of the functions $\varphi_1, \ldots, \varphi_n$, as functions of the parameters $x_1, \ldots, x_m$. This arbitrary choice will, of course, fix the form of the functions $\varphi_{rs} = \partial\varphi_r/\partial x_s$ and therefore will determine the value of the integral $\mathfrak{L}$ given in Eq. (3.1.1). Now let us change the φ's a little bit; for the function φ_r suppose the change to be expressed by the term $\epsilon_r\eta_r$, where η_r is an arbitrary function of the parameters and ϵ is a small quantity, independent of the parameters. The shorthand notation $\delta\varphi_r$ is often used instead of $\epsilon_r\eta_r$, where $\delta\varphi$ is considered to be an arbitrary small "variation" of the function φ. This modification of the φ's will also result in a change in the components φ_{rs} of the gradients. These are related to the changes in the φ's, for $\partial\epsilon_r\eta_r/\partial x_s = \epsilon_r\eta_{rs}$. In the shorthand variational notation, this relation is represented as $\delta\varphi_{rs} = \partial\delta\varphi_r/\partial x_s$.

By using a Taylor's series expansion of L we can show that the first-order change in the integral $\mathfrak{L}$ due to the small changes of the φ's can be written

$$\delta\mathfrak{L} = \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} \sum_{r=1}^{n} \epsilon_r \left[\frac{\partial L}{\partial \varphi_r} \eta_r + \sum_{s=1}^{m} \frac{\partial L}{\partial \varphi_{rs}} \frac{\partial \eta_r}{\partial x_s} \right] dx_1 \cdots dx_m$$

We assume that the parameters are so chosen that the limits of integration can all be constants and that all the η's go to zero at these limits, which would be true, for instance, if the limits coincided with physical boundaries where certain boundary conditions are imposed on the φ's. This situation is usually the case, so we shall confine ourselves here to its study; the more general case, where limits are all varied, will be touched upon later.

The term $(\partial L/\partial\varphi_{rs})(\partial\eta_r/\partial x_s)$ can be integrated by parts over x_s, giving $\left[\frac{\partial L}{\partial \varphi_{rs}} \eta_r\right]_{a_s}^{b_s} - \int_{a_s}^{b_s} \frac{\partial}{\partial x_s}\left(\frac{\partial L}{\partial \varphi_{rs}}\right) \eta_r \, dx_s$. The first term is zero, since $\eta_r = 0$ at a_s and b_s. Therefore the first-order variation of $\mathfrak{L}$ is

$$\delta\mathfrak{L} = \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} \sum_{r=1}^{n} \epsilon_r \left[\frac{\partial L}{\partial \varphi_r} - \sum_{s=1}^{m} \frac{\partial}{\partial x_s}\left(\frac{\partial L}{\partial \varphi_{rs}}\right) \right] \eta_r \, dx_1 \cdots dx_m \quad (3.1.2)$$

If $\delta\mathfrak{L}$ is not zero, $\mathfrak{L}$ cannot be a maximum or minimum. When $\delta\mathfrak{L}$ is zero, no matter what (small) values the ϵ_r's have, then the functional forms of all the φ's have been chosen so that $\mathfrak{L}$, as a function of the ϵ_r's, has *either a minimum or a maximum or a point of inflection* for $\epsilon_r = 0$. Usually we can tell from the physical situation which of these cases is true; if we are not sure, it is always possible to calculate the second-order terms in the ϵ_r's in the Taylor's series expansion of $\mathfrak{L}$ to see whether they are positive, negative, or zero. To save space from here on we shall use the terms "minimize" and "minimum" when we mean "minimized or maximized or minimaxed" and "minimum or maximum or point of inflection."

Therefore in order that $\mathfrak{L}$ have its extreme value (maximum or minimum), the functional form of the φ's must be chosen so that the coefficient of each of the ϵ_r's in the integral for $\delta\mathfrak{L}$ is zero. This results in a set of equations for the desired behavior of the φ's:

$$\sum_{s=1}^{m} \frac{\partial}{\partial x_s}\left(\frac{\partial L}{\partial \varphi_{rs}}\right) = \frac{\partial L}{\partial \varphi_r}; \quad r = 1, \ldots, n \quad (3.1.3)$$

where $\varphi_{rs} = \partial\varphi_r/\partial x_s$. These equations, which serve to determine the optimum functional form of the φ's, are called the *Euler equations*. We shall use these equations extensively later in this chapter.

Several general comments should be made concerning these results. In the first place, if the variational principle is to be of general validity, then $\mathfrak{L}$ should be an invariant, and the density L, or L divided by the scale factors coming into the element of integration, should be invariant to coordinate transformation of the parameters of integration. This will be of use later in finding new Lagrange densities.

A still more general comment is that the variational principle is generally useful in unifying a subject and consolidating a theory rather than in breaking ground for a new advance. It usually happens that the differential equations for a given phenomenon are worked out first, and only later is the Lagrange function found, from which the differential equations can be obtained. This is not to minimize the importance of finding the Lagrange density L, for it is of considerable utility to find what physical quantity must be minimized to obtain the differential equations for the phenomena, and the form of the variational equations often suggest fruitful analogies and generalizations.

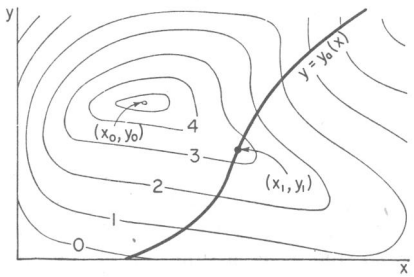

Fig. 3.1 Maximum point (x_0,y_0) for function $f(x,y,)$ represented by contour lines 0, 1, 2, ... Maximum point $(x_1,y_1,)$ along line $y = y_a(x)$.

Auxiliary Conditions. In many cases the Lagrange integral is to be minimized (or maximized) subject to some one or more additional requirements further restricting the independent variables and parameters. In this case we use the method of Lagrange multipliers to obtain the modified answer. Just how these multipliers work can best be shown by an example.

Suppose that the function $f(x,y)$ is to be maximized. If there are no auxiliary conditions, we solve the two equations

$$\partial f/\partial x = 0; \quad \partial f/\partial y = 0 \tag{3.1.4}$$

simultaneously. The resulting pair (or pairs) of values of x and y, (x_0,y_0), specify the point (or points) at which f has a maximum, minimum, or saddle point, and the value $f(x_0,y_0)$ is the value of f at this maximum or minimum. Here the function is of two variables, and the two equations (3.1.4) are needed to obtain a pair of values (x_0,y_0) for the maximum or minimum. A typical case is pictured in Fig. 3.1, where the function f is depicted in terms of contour lines.

But suppose that we wish to find the maximum of $f(x,y)$ along the line

given by the *auxiliary equation* $y = y_a(x)$. This line does not usually run through the point (x_0,y_0), so the solution cannot be the same. There may be one or more points along the line, however, where $f(x,y)$ has a maximum (or minimum) value, such as the point (x_1,y_1) shown in Fig. 3.1. This may be computed by inserting the expression for y in terms of x into the form for f, which gives the value of f along the line as a function the *single* parameter x. We then differentiate with respect to x to find the maximum value,

$$\frac{d}{dx} f(x,y_a(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{d}{dx}[y_a(x)] = 0 \tag{3.1.5}$$

The position of the maximum is then the solution x_1 of this equation and the related value $y_1 = y_a(x_1)$.

However, we can solve this same problem by a method which at first sight appears to be different from and more complicated than the one resulting in Eq. (3.1.5). Suppose the auxiliary equation is $g(x,y) = 0$. We first introduce a third unknown, λ, and then maximize the new function $f + \lambda g$, subject to the relation $g = 0$. In other words we are to solve the three equations

$$(\partial f/\partial x) + \lambda(\partial g/\partial x) = 0; \quad (\partial f/\partial y) + \lambda(\partial g/\partial y) = 0; \quad g = 0 \tag{3.1.6}$$

simultaneously to determine the proper values for x, y, *and* λ.

It is not immediately apparent that the solution of Eqs. (3.1.6) is identical with the solution of Eq. (3.1.5), but the connection becomes clearer when we write the auxiliary equation $g(x,y) = 0$ in the form used above, $y_a(x) - y = 0$. Then the first two of Eqs. (3.1.6) are

$$(\partial f/\partial x) + \lambda(dy_a/dx) = 0; \quad (\partial f/\partial y) - \lambda = 0$$

Substituting the second of these into the first gives us

$$(\partial f/\partial x) + (dy_a/dx)(\partial f/\partial y) = 0$$

which is just Eq. (3.1.5). Therefore in this simple case the method of Lagrange multipliers gives us the same result as the straightforward method. This is true in general. In this simple case it appears to be a more cumbersome method than the use of Eq. (3.1.5), but in more complex cases it turns out to be an easier method.

As applied to the variational integral of Eq. (3.1.1) the method of Lagrange multipliers can be stated as follows: Suppose that the Lagrange density is $L(\varphi_r,\varphi_{rs},x_s)$, $(s = 1, 2, \ldots, m)$, $(r = 1, 2, \ldots, n)$, and the auxiliary equations are

$$\int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} G_t\left(\varphi, \frac{\partial\varphi}{\partial x}, x\right) dx_1 \cdots dx_m = C_t \tag{3.1.7}$$

where $t = 1, 2, \ldots, k$ $(k < m)$ and where the C's are constants.

Then the modified variational integral is

$$\mathcal{L}' = \int_{a_1}^{b_1} \cdots \int_{a_m}^{b_m} L'\left(\varphi, \frac{\partial\varphi}{\partial x}, x\right) dx_1 \cdots dx_m$$

where

$$L' = L(\varphi_r, \varphi_{rs}, x_s) + \sum_{t=1}^{k} \lambda_t G_t(\varphi_r, \varphi_{rs} x_s) \tag{3.1.8}$$

Then the m new Euler equations,

$$\sum_{s=1}^{n} \frac{\partial}{\partial x_s}\left(\frac{\partial L'}{\partial \varphi_{rs}}\right) = \frac{\partial L'}{\partial \varphi_s} \tag{3.1.9}$$

plus the k equations (3.1.7) serve to determine the φ's as well as the values of the λ's. In this case the Lagrange multiplier method is definitely the easier.

3.2 *Hamilton's Principle and Classical Dynamics*

In classical dynamics the parameter is time t and the functions φ in the Lagrange function are the coordinates q which specify the configuration of the system at each time. If the system has n degrees of freedom, we can choose n independent q's $(q_1, \ldots, q_n)$ which will completely specify the configuration; the corresponding velocities will be $\dot{q}_r = dq_r/dt$. No matter how the q's are chosen, the kinetic energy of an inertial system always turns out to be a quadratic function of the $\dot{q}$'s:

$$T = \tfrac{1}{2} \sum_{r,s} a_{rs} \dot{q}_r \dot{q}_s \tag{3.2.1}$$

where the a's may be functions of the q's. If the system is *conservative* (*i.e.*, has a total mechanical energy which is constant in time), then the external force on the system can be represented in terms of the gradient of a scalar potential function V, so that the equivalent force along the q_r coordinate is

$$F_r = -(\partial V/\partial q_r) \tag{3.2.2}$$

The potential energy may depend on time explicitly, but it is not a function of the $\dot{q}$'s.

When the system is conservative, the variational principle determining the equations of motion, called *Hamilton's principle*, uses the *kinetic potential* (see page 229) $T - V$ as the Lagrange function and is

$$\delta \int_{t_0}^{t_1} (T - V)\, dt = 0 \tag{3.2.3}$$

This states that for any actual motion of the system, under the influence of the conservative forces, when it is started according to any reasonable initial conditions, the system will move so that the time average of the difference between kinetic and potential energies will be a minimum (or in a few cases, a maximum).

Lagrange's Equations. The Euler equations for the coordinates for this case,

$$\frac{d}{dt}\left[\frac{\partial(T - V)}{\partial \dot{q}_r}\right] = \frac{\partial(T - V)}{\partial q_r}; \quad r = 1, 2, \ldots, n \tag{3.2.4}$$

are called *Lagrange's equations of motion* for the system. The left-hand terms represent the accelerations of the system, and the right-hand terms are the corresponding forces, derived from the potential energy V, plus the "kinetic forces" (such as centrifugal force) due to the motions. When the forces are not conservative, so that there is no potential energy, the variational principle is

$$\int_{t_0}^{t_1} \left[\delta T + \sum_{r=1}^{n} F_r\, \delta q_r\right] dt = 0$$

and Lagrange's equations are

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_r}\right) = \frac{\partial T}{\partial q_r} - F_r \tag{3.2.5}$$

Lagrange's equations are among the most generally useful equations of classical dynamics.

The kinetic and potential energies are scalar invariants to coordinate transformations, so that they can be expressed in terms of any complete set of coordinates, and Lagrange's equations will have the same form in each case. In each case the quantity

$$\frac{\partial}{\partial \dot{q}_r}(T - V) = \frac{\partial T}{\partial \dot{q}_r} = p_r$$

is called the *momentum* for the rth coordinate. Lagrange's equations can be written

$$\frac{dp_r}{dt} - \frac{\partial T}{\partial q_r} = \begin{cases} -(\partial V/\partial q_r); & \text{if the system is conservative} \\ F_r; & \text{in general} \end{cases}$$

and if we use rectangular coordinates for the q's, so that T does not depend on the q's but only on the $\dot{q}$'s, the equations reduce to the familiar Newton's equations

$$(d/dt)(\text{momentum}) = \text{force}$$

Consequently Hamilton's principle represents, in simple invariant form, all the equations of classical dynamics.