CHAPTER 1

CLASSICAL FIELDS

1–1. PARTICLES AND FIELDS

Nonrelativistic quantum mechanics, developed in the years from 1923 to 1926, provides a unified and logically consistent picture of numerous phenomena in the atomic and molecular domain. Following P.A.M. Dirac, we might be tempted to assert: "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are completely known."

There are, however, basically two reasons for believing that the description of physical phenomena based on nonrelativistic quantum mechanics is incomplete. First, since nonrelativistic quantum mechanics is formulated in such a way as to yield the nonrelativistic energy-momentum relation in the classical limit, it is incapable of accounting for the fine structure of a hydrogen-like atom. (This problem was treated earlier by A. Sommerfeld, who used a relativistic generalization of N. Bohr's atomic model.) In general, nonrelativistic quantum mechanics makes no prediction about the dynamical behavior of particles moving at relativistic velocities. This defect was amended by the relativistic theory of electrons developed by Dirac in 1928, which will be discussed in Chapter 3. Second, and what is more serious, nonrelativistic quantum mechanics is essentially a single-particle theory in which the probability density for finding a given particle integrated over all space is unity at all times. Thus it is not constructed to describe phenomena such as nuclear beta decay in which an electron and an antineutrino are created as the neutron becomes a proton or to describe even a simpler process in which an excited atom returns to its ground state by "spontaneously" emitting a single photon in the absence of any external field. Indeed, it is no accident that many of the most creative theoretical physicists in the past forty years have spent their main efforts on attempts to understand physical phenomena in which various particles are created or annihilated. The major part of this book is devoted to the progress physicists have made along these lines since the historic 1927 paper of Dirac entitled "The Quantum Theory of the Emission and Absorption of Radiation" opened up a new subject called the *quantum theory of fields*.

The concept of a field was originally introduced in classical physics to account for the interaction between two bodies separated by a finite distance. In classical physics the electric field $\mathbf{E}(\mathbf{x}, t)$, for instance, is a three-component function defined at each space-time point, and the interaction between two charged bodies, 1 and 2, is to be viewed as the interaction of body 2 with the electric field created by body 1. In the quantum theory, however, the field concept acquires a new dimen-

sion. As originally formulated in the late 1920's and the early 1930's, the basic idea of quantum field theory is that we associate *particles* with fields such as the electromagnetic field. To put it more precisely, quantum-mechanical excitations of a field appear as particles of definite mass and spin, a notion we shall illustrate in Section 2-2, where the connection between the transverse electromagnetic field and photons is discussed in detail.

Even before the advent of postwar calculational techniques which enabled us to compute quantities such as the $2s$–$2p_{1/2}$ separation of the hydrogen atom to an accuracy of one part in 10^8, there had been a number of brilliant successes of the quantum theory of fields. First, as we shall discuss in Chapter 2, the quantum theory of radiation developed by Dirac and others provides quantitative understandings of a wide class of phenomena in which real photons are emitted or absorbed. Second, the requirements imposed by quantum field theory, when combined with other general principles such as Lorentz invariance and the probabilistic interpretation of state vectors, severely restrict the class of particles that are permitted to exist in nature. In particular, we may cite the following two rules derivable from *relativistic* quantum field theory:

a) For every charged particle there must exist an antiparticle with opposite charge and with the same mass and lifetime.
b) The particles that occur in nature must obey the spin-statistics theorem (first proved by W. Pauli in 1940) which states that half-integer spin particles (e.g., electron, proton, Λ-hyperon) must obey Fermi-Dirac statistics, whereas integer spin particles (e.g., photon, π-meson, K-meson) must obey Bose-Einstein statistics.

Empirically there is no known exception to these rules. Third, the existence of a nonelectromagnetic interaction between two nucleons at short but finite distances prompts us to infer that a field is responsible for nuclear forces; this, in turn, implies the existence of massive particles associated with the field, a point first emphasized by H. Yukawa in 1935. As is well known, the desired particles, now known as π-mesons or pions, were found experimentally twelve years after the theoretical prediction of their existence.

These considerations appear to indicate that the idea of associating particles with fields and, conversely, fields with particles is not entirely wrong. There are, however, difficulties with the present form of quantum field theory which must be overcome in the future. First, as we shall show in the last section of Chapter 4, despite the striking success of postwar quantum electrodynamics in calculating various observable effects, the "unobservable" modifications in the mass and charge of the electron due to the emission and reabsorption of a virtual photon turn out to diverge logarithmically with the frequency of the virtual photon. Second, the idea of associating a field with each "particle" observed in nature becomes ridiculous and distasteful when we consider the realm of strong interactions where many different kinds of "particles" are known to interact with one another: we know from experiment that nearly 100 "particles" or "resonances" participate in the physics of strong interactions. This difficulty became particularly acute in 1961–1964 when a successful classification scheme of strongly interacting

particles was formulated which groups together into a single "family" highly unstable "particles" (lifetimes 10^{-23} sec, often called strong interaction resonances) and moderately metastable particles (lifetimes 10^{-10} sec).‡ Yet, despite these difficulties, it is almost certain that there are many elements in present-day quantum field theory which are likely to survive, say, one hundred years from now.

Before we study quantized fields, we will study classical fields. In part this decision is motivated by the historical fact that prior to the development of quantum electrodynamics there was the classical electrodynamics of Maxwell which, among other things, successfully predicted the existence of Hertzian electromagnetic waves. This chapter is primarily concerned with the elements of *classical* field theory needed for the understanding of *quantized* fields. As a preliminary to the study of quantization we are particularly interested in the dynamical properties of classical fields. For this reason we will follow an approach analogous to Hamilton's formulation of Lagrangian mechanics.

1-2. DISCRETE AND CONTINUOUS MECHANICAL SYSTEMS

The dynamical behavior of a single particle, or more precisely, a mass point in classical mechanics, can be inferred from Lagrange's equation of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \tag{1.1}$$

which is derivable from Hamilton's variational principle

$$\delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i)\, dt = 0. \tag{1.2}$$

The Lagrangian L (assumed here not to depend explicitly on time) is given by the difference of the kinetic energy T and the potential energy V,

$$L = T - V, \tag{1.3}$$

and the variation in (1.2) is to be taken over an arbitrary path $q_i(t)$ such that δq_i vanishes at t_1 and t_2. The Hamiltonian of the system is

$$H = \sum_i p_i \dot{q}_i - L, \tag{1.4}$$

where the momentum p_i, canonical conjugate to q_i, is given by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{1.5}$$

‡In fact the one-to-one correspondence between a "field" and a "particle" appears to be lost in a more modern formulation of the field theory of strong interactions as many (if not all) of the so-called "elementary" particles may well be regarded as bound (or resonant) states of each other. The distinction between fundamental particles and composite states, however, is much more clear-cut in the realm of the electromagnetic interactions among electrons, muons, and photons. As an example, in Section 4-4 we shall calculate the lifetime of the ground state of positronium without introducing a field corresponding to the positronium.

These considerations can be generalized to a system with many particles. As a concrete example, let us consider a collection of N particles connected with identical springs of force constant k and aligned in one dimension, as shown in Fig. 1–1.‡ By calling η_i the displacement of the ith particle from its equilibrium position we write the Lagrangian L as follows:

$$L = \tfrac{1}{2}\sum_i^N [m\dot{\eta}_i^2 - k(\eta_{i+1} - \eta_i)^2]$$
$$= \sum_i^N a\,\frac{1}{2}\left[\frac{m}{a}\dot{\eta}_i^2 - ka\left(\frac{\eta_{i+1} - \eta_i}{a}\right)^2\right]$$
$$= \sum_i^N a\mathscr{L}_i, \qquad (1.6)$$

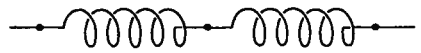

Fig. 1–1. Particles connected with identical springs.

where a is the separation distance between the equilibrium positions of two neighboring particles and $\mathscr{L}_i$ is the linear Lagrangian density, i.e. the Lagrangian density per unit length.

We can pass from the above discrete mechanical system to a continuous mechanical system as the number of degrees of freedom becomes infinite in such a way that the separation distance becomes infinitesimal:

$$a \to dx, \qquad \frac{m}{a} \to \mu = \text{linear mass density},$$
$$\frac{\eta_{i+1} - \eta_i}{a} \to \frac{\partial \eta}{\partial x}, \qquad ka \to Y = \text{Young's modulus}. \qquad (1.7)$$

We now have

$$L = \int \mathscr{L}\, dx, \qquad (1.8)$$

where

$$\mathscr{L} = \tfrac{1}{2}\left[\mu\dot{\eta}^2 - Y\left(\frac{\partial \eta}{\partial x}\right)^2\right]. \qquad (1.9)$$

We note that η itself has become a function of the continuous parameters x and t. Yet in the Lagrangian formalism η should be treated like a generalized "coordinate" just as q_i in L of Eq. (1.2).

In formulating the variational principle in the continuous case we consider

$$\delta\int_{t_1}^{t_2} L\, dt = \delta\int_{t_1}^{t_2} dt \int dx\, \mathscr{L}\left(\eta, \dot{\eta}, \frac{\partial \eta}{\partial x}\right). \qquad (1.10)$$

The variation on η is assumed to vanish at t_1 and t_2 and also at the extremities of the space integration. (In field theory this latter requirement is not stated explicitly since we are usually considering a field which goes to zero sufficiently rapidly at infinity.) Otherwise the nature of the variation is completely arbitrary. The variational integral becomes

$$\delta\int L\, dt = \int dt \int dx \left\{\frac{\partial \mathscr{L}}{\partial \eta}\delta\eta + \frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial x)}\delta\left(\frac{\partial \eta}{\partial x}\right) + \frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial t)}\delta\left(\frac{\partial \eta}{\partial t}\right)\right\}$$
$$= \int dt \int dx \left\{\frac{\partial \mathscr{L}}{\partial \eta}\delta\eta - \frac{\partial}{\partial x}\left(\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial x)}\right)\delta\eta - \frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial t)}\delta\eta\right\}, \qquad (1.11)$$

‡This problem is treated in greater detail in Goldstein (1951), Chapter 11.

where the integrations by parts of the last two terms can be justified since $\delta\eta$ vanishes at the end points of the space and time intervals. If (1.11) is to vanish for any arbitrary variation satisfying the above requirements, we must have

$$\frac{\partial}{\partial x}\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial x)} + \frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial t)} - \frac{\partial \mathscr{L}}{\partial \eta} = 0. \qquad (1.12)$$

This is called the Euler-Lagrange equation.‡ In our particular example (1.9), Eq. (1.12) becomes

$$Y\frac{\partial^2\eta}{\partial x^2} - \mu\frac{\partial^2\eta}{\partial t^2} = 0. \qquad (1.13)$$

This is to be identified with the wave equation for the one-dimensional propagation of a disturbance with velocity $\sqrt{Y/\mu}$. We can define the Hamiltonian density $\mathscr{H}$ in analogy with (1.4) as

$$\mathscr{H} = \dot{\eta}\frac{\partial \mathscr{L}}{\partial \dot{\eta}} - \mathscr{L}$$
$$= \tfrac{1}{2}\mu\dot{\eta}^2 + \tfrac{1}{2}Y\left(\frac{\partial \eta}{\partial x}\right)^2; \qquad (1.14)$$

$\partial\mathscr{L}/\partial\dot{\eta}$ is called the canonical momentum conjugate to η, and is often denoted by π. The two terms in (1.14) can be identified respectively with the kinetic and potential energy densities.

1–3. CLASSICAL SCALAR FIELDS

Covariant notation. The arguments of the preceding section can readily be generalized to three space dimensions. Consider a field which is assumed to be a real function defined at each space-time point, $\mathbf{x}$, t; $\mathscr{L}$ now depends on ϕ, $\partial\phi/\partial x_k$ ($k = 1, 2, 3$), and $\partial\phi/\partial t$. The Euler-Lagrange equation reads

$$\sum_{k=1}^{3}\frac{\partial}{\partial x_k}\frac{\partial \mathscr{L}}{\partial(\partial\phi/\partial x_k)} + \frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial(\partial\phi/\partial t)} - \frac{\partial \mathscr{L}}{\partial \phi} = 0. \qquad (1.15)$$

We wish to write (1.15) in a relativistically covariant form, but first let us recall some properties of Lorentz transformations. We introduce a four-vector notation in which the four-vector b_μ with $\mu = 1, 2, 3, 4$ stands for

$$b_\mu = (b_1, b_2, b_3, b_4) = (\mathbf{b}, ib_0), \qquad (1.16)$$

where b_1, b_2, and b_3 are real, and $b_4 = ib_0$ is purely imaginary. In general, the Greek indices μ, ν, λ, etc., run from 1 to 4, whereas the italic indices i, j, k, etc.,

‡In the literature this equation is sometimes writen in the form

$$\frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial t)} - \frac{\delta \mathscr{L}}{\delta \eta} = 0,$$

where $\delta\mathscr{L}/\delta\eta$ is called the functional derivative of $\mathscr{L}$ with respect to η. This version is not recommended since (a) it obscures the dependence of $\mathscr{L}$ on the space coordinate, and (b) it singles out time, which is against the spirit of the covariant approach (to be discussed in the next section).

run from 1 to 3. The coordinate vector x_μ is given by

$$\begin{aligned} x_\mu &= (x_1, x_2, x_3, x_4) \\ &= (\mathbf{x}, ict). \end{aligned} \tag{1.17}$$

The symbols x, y, and z may also be used in place of x_1, x_2, and x_3. Under a Lorentz transformation, we have

$$x'_\mu = a_{\mu\nu} x_\nu, \tag{1.18}$$

where the $a_{\mu\nu}$ satisfy

$$a_{\mu\nu} a_{\mu\lambda} = \delta_{\nu\lambda}, \qquad (a^{-1})_{\mu\nu} = a_{\nu\mu}. \tag{1.19}$$

Hence

$$x_\mu = (a^{-1})_{\mu\nu} x'_\nu = a_{\nu\mu} x'_\nu \tag{1.20}$$

when x' and x are related by (1.18). The matrix elements a_{ij}, a_{44} are purely real, whereas a_{j4} and a_{4j} are purely imaginary. A *four-vector*, by definition, transforms in the same way as x_μ under Lorentz transformations. Because of (1.20) we have

$$\frac{\partial}{\partial x'_\mu} = \frac{\partial x_\nu}{\partial x'_\mu} \frac{\partial}{\partial x_\nu} = a_{\mu\nu} \frac{\partial}{\partial x_\nu}; \tag{1.21}$$

so the four-gradient $\partial/\partial x_\mu$ is a four-vector. The scalar product $b \cdot c$ is defined by

$$\begin{aligned} b \cdot c &= b_\mu c_\mu = \sum_{j=1}^{3} b_j c_j + b_4 c_4 \\ &= \mathbf{b} \cdot \mathbf{c} - b_0 c_0. \end{aligned} \tag{1.22}$$

It is unchanged under Lorentz transformations, since

$$\begin{aligned} b' \cdot c' &= a_{\mu\nu} b_\nu a_{\mu\lambda} c_\lambda = \delta_{\nu\lambda} b_\nu c_\lambda \\ &= b \cdot c. \end{aligned} \tag{1.23}$$

A tensor of second rank, $t_{\mu\nu}$, transforms as

$$t'_{\mu\nu} = a_{\mu\lambda} a_{\nu\sigma} t_{\lambda\sigma}. \tag{1.24}$$

Generalizations to tensors of higher rank are straightforward. Note that we make no distinction between a covariant and a contravariant vector, nor do we define the metric tensor $g_{\mu\nu}$. These complications are absolutely unnecessary in the *special* theory of relativity. (It is regrettable that many textbook writers do not emphasize this elementary point.)

Equation (1.15) can now be written as

$$\frac{\partial}{\partial x_\mu} \left[\frac{\partial \mathscr{L}}{\partial(\partial\phi/\partial x_\mu)} \right] - \frac{\partial \mathscr{L}}{\partial \phi} = 0. \tag{1.25}$$

It is seen that the field equation derivable from the Lagrangian density $\mathscr{L}$ is covariant (i.e., the equation "looks the same" in all Lorentz frames) if the Lagrangian density $\mathscr{L}$ is chosen to be a relativistically scalar density. This is an important point because the relativistic invariance of $\mathscr{L}$ is so restrictive that it can be used as a guiding principle for "deriving" a covariant wave equation.

Neutral scalar field. As an illustration let $\phi(x)$ be a scalar field which, by definition, transforms like

$$\phi'(x') = \phi(x), \tag{1.26}$$

under a Lorentz transformation, where ϕ' is the functional form of the field in the primed system. Now the dependence of $\mathscr{L}$ on space-time coordinates is only through the field and its first derivatives, and x_μ cannot appear explicitly in $\mathscr{L}$. This means that $\partial\phi/\partial x_\mu$ is the only four-vector at our disposal; when it appears in $\mathscr{L}$ it must be contracted with itself. Moreover, if we are interested in obtaining a linear wave equation, $\mathscr{L}$ must be a quadratic function of ϕ and $\partial\phi/\partial x_\mu$. A possible candidate for $\mathscr{L}$ consistent with the above requirements is

$$\mathscr{L} = -\frac{1}{2} \left(\frac{\partial \phi}{\partial x_\mu} \frac{\partial \phi}{\partial x_\mu} + \mu^2 \phi^2 \right). \tag{1.27}$$

From the Euler-Lagrange equation (1.25) we obtain

$$-\frac{1}{2} \frac{\partial}{\partial x_\mu} \left(2 \frac{\partial \phi}{\partial x_\mu} \right) + \mu^2 \phi = 0, \tag{1.28}$$

or

$$\Box \phi - \mu^2 \phi = 0, \tag{1.29}$$

where

$$\Box = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}. \tag{1.30}$$

The wave equation (1.29) is called the Klein-Gordon equation. It was considered in the middle 1920's by E. Schrödinger, as well as by O. Klein and W. Gordon, as a candidate for the *relativistic* analog of the *nonrelativistic* Schrödinger wave equation for a free particle. The similarity of (1.29) to the relativistic energy momentum relation for a free particle of mass m,

$$E^2 - |\mathbf{p}|^2 c^2 = m^2 c^4, \tag{1.31}$$

becomes apparent as we consider heuristic substitutions:

$$E \to i\hbar \frac{\partial}{\partial t}, \qquad p_k \to -i\hbar \frac{\partial}{\partial x_k}. \tag{1.32}$$

The parameter μ in (1.29) has the dimension of inverse length, and, using (1.32), we may make the identification

$$\mu = mc/\hbar. \tag{1.33}$$

Numerically $1/\mu$ is 1.41×10^{-13} cm for a particle of mass 140 MeV/c^2 (corresponding to the mass of the charged pion).

Yukawa potential. So far we have been concerned with a field in the absence of any source. Such a field is often called a *free field*. The interaction of ϕ with a source can easily be incorporated into the Lagrangian formalism by adding

$$\mathscr{L}_{\text{int}} = -\phi\rho, \tag{1.34}$$

to (1.27), where ρ is the source density, which is, in general, a function of space-time coordinates. The field equation now becomes

$$\Box \phi - \mu^2 \phi = \rho. \tag{1.35}$$