1
Introduction

1.1 Overview

It is a familiar fact that when a field theory is treated quantum mechanically the wave solutions of the classical theory lead to elementary quanta that have a natural interpretation as particles in the quantum theory. This suggests a one-to-one correspondence between fields and particle species and is the basis for the standard applications of perturbative quantum field theory.

However, many classical field theories have solutions that are already particle-like at the classical level. These are characterized by an energy density that is localized in space and that does not dissipate over time. It is natural to ask whether these "solitons", as they are called, have counterparts in the quantum version of the theory. If so, they would presumably be a new species of particle, quite distinct from the "elementary" particle associated with the wave solutions of the free field theory.

It is instructive to compare the classical size of the soliton with the Compton wavelength that it would have in the quantum theory. If the elementary particles of the theory have masses of order m and a characteristic coupling of order g, one typically finds that the soliton has a classical energy

$$E_{\text{classical}} \sim \frac{m}{g} \tag{1.1}$$

and a characteristic spatial size $\ell_{\text{soliton}} \sim 1/m$. Hence,

$$\lambda_{\text{Compton}} \sim \frac{1}{E_{\text{classical}}} \sim g\,\ell_{\text{soliton}}. \tag{1.2}$$

(I am using units with $\hbar = 1$.) If the coupling is weak, the Compton wavelength is much less than the classical size, and so we might expect the soliton to survive, perhaps with slight modifications, after quantization.

A possible objection is the stark contrast between the smooth profile of the classical solution and the fuzziness of quantum field theory. It is certainly true that the quantum fluctuations of the field are large, even divergent, when the field is measured at very short distances. However, these fluctuations are reduced when the field is averaged over a larger smearing distance. We will see that the same weak-coupling regime that gives $\ell_{\text{soliton}} \gg \lambda_{\text{Compton}}$ also guarantees the existence of a smearing distance that is both large enough to suppress the quantum fluctuations and small enough that the classical field profile is still evident.

The inverse dependence on the coupling implies that in this weak-coupling regime the soliton mass is large, tending toward infinity as the coupling goes to zero. This explains why the effects of the soliton are not seen in perturbation theory. Nevertheless, once the classical solution is known, perturbative methods can be used to quantize the fields about the soliton and to demonstrate that there is indeed a corresponding one-particle state in the quantum theory. Furthermore, the quantum corrections to the classical energy are calculable and give a mass of the form

$$M_{\text{quantum}} = E_{\text{classical}} \left(1 + c_1 g + c_2 g^2 + \cdots\right). \tag{1.3}$$

What about the strong-coupling regime? Even though the soliton may still be a solution of the classical field equations, the perturbative analysis of the quantum theory breaks down here, and the arguments for a quantum counterpart to the soliton are no longer so clear-cut. However, a new and striking phenomenon may now come into play. There are examples of theories—a particularly well-known pair being the sine-Gordon and massive Thirring models—that are related by a duality that maps the weak-coupling regime of one onto the strong-coupling regime of the other. The sine-Gordon soliton states correspond to elementary particle states of the massive Thirring model, while the elementary particle of the sine-Gordon model becomes a massive Thirring bound state. One must conclude that there is no intrinsic difference between an elementary particle and a soliton. The distinction between them is simply that one viewpoint or the other is more convenient for calculation in a particular coupling regime.

Although we live in a world with three spatial dimensions (and perhaps some additional hidden ones), it can be instructive to consider solitons in fewer dimensions. The analysis of these toy models is often more tractable and helps elucidate issues of principle. Their solutions can also be trivially extended to higher dimensions, where they acquire new physical significance. A particle-like soliton in one dimension can be interpreted as a planar solution in three dimensions, corresponding to a domain wall. Similarly, a two-dimensional particle-like soliton becomes a line solution, or string, in three dimensions.

One can also consider solitons in more than three spatial dimensions. Of particular interest are those in four dimensions. These could be viewed as particles in a hypothetical world with four spatial dimensions. Alternatively, and more importantly, they can be interpreted as solutions in a Euclideanized version of our four-dimensional spacetime. Such Euclidean solutions, or instantons, have no obvious physical significance in a classical context. However, they become meaningful quantum mechanically because wavefunctions extend into classically forbidden regions where the potential energy is greater than the total energy. Roughly speaking, one can view this as implying a negative kinetic energy, corresponding to evolution in a Euclidean spacetime with imaginary time. A well-known consequence is that quantum systems can tunnel though potential energy barriers to effect transitions that would be classically forbidden. This leads to important and unexpected nonperturbative effects in gauge theories, with magnitudes that are determined by the action of the relevant instanton. A further result of tunneling processes in field theory is the decay of metastable vacua by bubble nucleation, a process of considerable importance for cosmology. The Euclidean solutions that govern such bubble nucleation are known as bounces.

Finally, a note on terminology. I follow the practice in high energy physics of using the term soliton for any localized classical solution that does not dissipate over time. However, the reader should be aware that some other fields use a more restrictive definition, with the term only used for solutions, arising in integrable systems, that emerge from scattering processes without deformation or loss of energy.

1.2 Conventions

Metric and indices

For the spacetime metric I use the "mostly minus" convention, with the metric $\eta_{\mu\nu} = \mathrm{diag}\,(1, -1, -1, -1)$ in flat four-dimensional spacetime. Coordinates are defined by

$$x^\mu = (t, x, y, z) = (t, \mathbf{x}) \tag{1.4}$$

so that

$$\partial_\mu = (\partial/\partial t, \boldsymbol{\nabla}). \tag{1.5}$$

Lorentzian spacetime indices are denoted by Greek letters and summation over repeated indices, one upper and one lower, is to be understood. Purely spatial indices are denoted by Latin letters, generally from the middle of the alphabet; summation over repeated indices (possibly both upper or both lower) is also to be understood. Euclidean spacetime indices are denoted by Latin letters.

The antisymmetric tensor in any dimension is defined to be unity when all of its indices are upper and in numerical order. Thus, $\epsilon^{123} = \epsilon^{0123} = \epsilon^{1234} = 1$.

Dirac matrices

The Dirac matrices in four-dimensional Lorentzian spacetime obey

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{1.6}$$

Of these, γ^0 is Hermitian, while the remaining three are anti-Hermitian. The matrix

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{1.7}$$

is Hermitian and obeys $(\gamma^5)^2 = I$.

Units

I use natural units with c, $\hbar$, and Boltzmann's constant k_B all equal to unity.

Gauge fields

Conventions associated with gauge fields vary within the soliton and instanton literature. Those used in this book are described below.

The electromagnetic potential is

$$A^\mu = (\Phi, \mathbf{A}) \tag{1.8}$$

where Φ and $\mathbf{A}$ are the usual scalar and vector potentials. The field strength tensor is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.9}$$

so that, e.g., $F_{12} = F^{12} = -B_z$ and $F_{03} = F^{30} = E_z$. The covariant derivative of a complex field carrying electromagnetic [or any other U(1)] charge q is given by

$$D_\mu \phi = (\partial_\mu + iqA_\mu)\phi. \tag{1.10}$$

The Lagrangian is then invariant under U(1) gauge transformations of the form

$$\begin{aligned} \phi &\to e^{iq\Lambda(x)}\phi, \\ A_\mu &\to A_\mu - \partial_\mu \Lambda(x). \end{aligned} \tag{1.11}$$

In non-Abelian gauge theories the gauge field is written as a Hermitian element of the Lie algebra

$$A_\mu = A_\mu^a T^a, \tag{1.12}$$

where the Hermitian generators T^a are normalized so that

$$\operatorname{tr} T^a T^b = \frac{1}{2}\delta^{ab}. \tag{1.13}$$

They obey

$$[T^a, T^b] = i f_{abc} T^c, \tag{1.14}$$

with the structure constants f_{abc} being totally antisymmetric. This corresponds to the standard normalization for the fundamental representation of SU(2), with the generators being $\sigma^a/2$, where the σ^a are the Pauli matrices. The field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu], \tag{1.15}$$

with components

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{abc} A_\mu^b A_\nu^c. \tag{1.16}$$

A matter field ϕ can be written as a column vector transforming under an irreducible representation of the gauge group. Its covariant derivative is

$$D_\mu \phi = \partial_\mu \phi - igA_\mu \phi. \tag{1.17}$$

With components written out explicitly, this is

$$(D_\mu \phi)_j = \partial_\mu \phi_j - igA_\mu^a (t^a)_{jk} \phi_k \tag{1.18}$$

with i, j, and k running from 1 to N and $(t^a)_{jk}$ denoting the appropriate representation of the generators.

Under a non-Abelian gauge transformation $U(x)$, the various quantities above transform as

$$\begin{aligned} A_\mu &\longrightarrow UA_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1} = UA_\mu U^{-1} + \frac{i}{g}U\,\partial_\mu U^{-1}, \\ F_{\mu\nu} &\longrightarrow UF_{\mu\nu}U^{-1}, \\ \phi &\longrightarrow \mathcal{U}\phi. \end{aligned} \tag{1.19}$$

where $\mathcal{U}$ is the transformation written in the appropriate representation of the group. For an infinitesimal gauge transformation

$$U = e^{i\Lambda} \approx I + i\Lambda + \cdots \tag{1.20}$$

the change in the gauge potential is

$$\delta A_\mu = \frac{1}{g}\partial_\mu \Lambda - i[A_\mu, \Lambda] = \frac{1}{g}D_\mu \Lambda. \tag{1.21}$$

If the matter fields transform under the adjoint representation, an alternative notation is to write them as linear combinations of the generators,

$$\phi = \phi^a T^a \tag{1.22}$$

with

$$D_\mu \phi = \partial_\mu \phi - ig[A_\mu, \phi]. \tag{1.23}$$

In the special case of a triplet field in an SU(2) gauge theory (where $f_{abc} = \epsilon_{abc}$) I sometimes adopt the standard three-dimensional vector notation and write

$$\begin{aligned} D_\mu \boldsymbol{\phi} &= \partial_\mu \boldsymbol{\phi} + g\mathbf{A}_\mu \times \boldsymbol{\phi}, \\ \mathbf{F}_{\mu\nu} &= \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu + g\mathbf{A}_\mu \times \mathbf{A}_\nu. \end{aligned} \tag{1.24}$$

It is sometimes convenient to absorb the gauge coupling in the gauge field by a rescaling $A_\mu \to gA_\mu$. The Yang–Mills Lagrangian is then

$$\mathcal{L} = -\frac{1}{4g^2}F_{\mu\nu}^a F^{\mu\nu a} = -\frac{1}{2g^2}\operatorname{tr} F_{\mu\nu}F^{\mu\nu}. \tag{1.25}$$