Cambridge University Press
0521813638 - Classical Covariant Fields
Mark Burgess
Excerpt
More information

Part 1
Fields

1
Introduction

In contemporary field theory, the word *classical* is reserved for an analytical framework in which the local equations of motion provide a complete description of the evolution of the fields. Classical field theory is a differential expression of change in functions of space and time, which summarizes the state of a physical system entirely in terms of smooth fields. The differential (holonomic) structure of field theory, derived from the action principle, implies that field theories are microscopically reversible by design: differential changes experience no significant obstacles in a system and may be trivially undone. Yet, when summed macroscopically, in the context of an environment, such individually reversible changes lead to the well known irreversible behaviours of thermodynamics: the reversal of paths through an environmental landscape would require the full history of the route taken. Classical field theory thus forms a basis for both the microscopic and the macroscopic.

When applied to quantum mechanics, the classical framework is sometimes called the *first quantization*. The first quantization may be considered the first stage of a more complete theory, which goes on to deal with the issues of many-particle symmetries and interacting fields. Quantum mechanics is classical field theory with additional assumptions about measurement. The term *quantum mechanics* is used as a name for the specific theory of the Schrödinger equation, which one learns about in undergraduate studies, but it is also sometimes used for any fundamental description of physics, which employs the measurement axioms of Schrödinger quantum mechanics, i.e. where change is expressed in terms of fields and groups. In that sense, this book is also about quantum mechanics, though it does not consider the problem of measurement, and all of its subtlety.

In the so-called *quantum field theory*, or *second quantization*, fields are promoted from c-number functions to operators, acting upon an additional set of states, called Fock space. Fock space supplants Slater determinant combinatorics in the classical theory, and adds a discrete aspect to smooth field theory. It quantizes the allowed amplitudes of the normal modes of the field and gives excitations the same denumerable property that ensembles of particles have; i.e. it adds *quanta* to the fields, or indistinguishable, countable excitations, with varying numbers. Some authors refer to these quanta simply as 'particles'; however, they are not particles in the classical sense of localizable, pointlike objects. Moreover, whereas particles are separate entities, quanta are excitations, spawned from a single entity: the quantum field. The second-quantized theory naturally incorporates the concept of a lowest possible energy state (the vacuum), which rescues the relativistic theory from negative energies and probabilities. Such an assumption must be added by hand in the classical theory. When one speaks about *quantum* field theory, one is therefore referring to this 'second quantization' in which the fields are dynamical operators, spawning indistinguishable quanta.

This book is not about quantum field theory, though one might occasionally imagine it is. It will mention the quantum theory of fields, only insofar as to hint at how it generalizes the classical theory of fields. It discusses statistical aspects of the classical field to the extent that classical Boltzmann statistical mechanics suffices to describe them, but does not delve into interactions or combinatorics. One should not be misled; books on quantum field theory generally begin with a dose of classical field theory, and many purely classical ideas have come to be confused with second-quantized ones. Only in the final chapter is the second-quantized framework outlined for comparison. This book is a summary of the core methodology, which underpins covariant field theory at the classical level. Rather than being a limitation, this avoidance of quantum field theory allows one to place a sharper focus on key issues of symmetry and causality which lie at the heart of all subsequent developments, and to dwell on the physical interpretation of formalism in a way which other treatments take for granted.

1.1 Fundamental and effective field theories

The main pursuit of theoretical physics, since quantum mechanics was first envisaged, has been to explore the maxim that the more microscopic a theory is, the more fundamental it is. In the 1960s and 1970s it became clear that this view was too simplistic. Physics is as much about *scale* as it is about constituent components. What is fundamental at one scale might be irrelevant to physics at another scale. For example, quark dynamics is not generally required to describe the motion of the planets. All one needs, in fact, is an effective theory of planets as point mass objects. their detailed structure is irrelevant to so many decimal places that it would be nonsense to attempt to include it in calculations. Planets are less elementary than quarks, but they are not less fundamental to the problem at hand.

The quantum theory of fields takes account of dynamical *correlations* between the field at different points in space and time. These correlations, called fluctuations or virtual processes, give rise to *quantum corrections* to the equations of motion for the fields. At first order, these can also be included in the classical theory. The corrections modify the form of the equations of motion and lead to *effective field equations* for the quantized system. At low energies, these look like classical field theories with renormalized coefficients. Indeed, this sometimes results in the confusion of statistical mechanics with the second quantization. Put another way, at a superficial level all field theories are approximately classical field theories, if one starts with the right coefficients. The reason for this is that all one needs to describe physical phenomena is a blend of two things: symmetry and causal time evolution. What troubles the second quantization is demonstrating the consistency of this point of view, given sometimes uncertain assumptions about space, time and the nature of fields.

This point has been made, for instance, by Wilson in the context of the renormalization group [139]; it was also made by Schwinger, in the early 1970s, who, disillusioned with the direction that field theory was taking, redefined his own interpretation of field theory called *source theory* [119], inspired by ideas from Shannon's mathematical theory of communication [123]. The thrust of source theory is the abstraction of irrelevant detail from calculations, and a reinforcement of the importance of causality and boundary conditions.

1.2 The continuum hypothesis

Even in classical field theory, there is a difference between particle and field descriptions of matter. This has nothing *a priori* to do with wave–particle duality in quantum mechanics. Rather, it is to do with scale.

In classical mechanics, individual pointlike particle trajectories are characterized in terms of 'canonical variables' $x(t)$ and $p(t)$, the position and momentum at time t. Underpinning this description is the assumption that matter can be described by particles whose important properties are localized at a special place at a special time. It is not even necessarily assumed that matter is made of particles, since the particle position might represent the centre of mass of an entire planet, for instance. The key point is that, in this case, the centre of mass is a localizable quantity, relevant to the dynamics.

In complex systems composed of many particles, it is impractical to take into account the behaviour of every single particle separately. Instead, one invokes the continuum hypothesis, which supposes that matter can be treated as a continuous substance with bulk properties at large enough scales. A system with a practically infinite number of point variables is thus reduced to the study of continuous functions or *effective fields*. Classically, continuum theory is a high-level or *long-wavelength* approximation to the particle theory, which blurs out the individual particles. Such a theory is called an *effective theory*.

In quantum mechanics, a continuous wavefunction determines the probability of measuring a discrete particle event. However, free elementary quantum particles cannot be localized to precise trajectories because of the uncertainty principle. This wavefunction-field is different from the continuum hypothesis of classical matter: it is a function which represents the state of the particle's quantum numbers, and the probability of its position. It is not just a smeared out approximation to a more detailed theory. The continuous, field nature is observed as the interference of matter waves in electron diffraction experiments, and single-particle events are measured by detectors. If the wavefunction is sharply localized in one place, the probability of measuring an event is very large, and one can argue that the particle has been identified as a bump in the field.

To summarize, a sufficient number of localizable particles can be viewed as an effective field, and conversely a particle can be viewed as a localized disturbance in an elementary field.

To envisage an elementary field as representing particles (not to be confused with quanta), one ends up with a picture of the particles as localized disturbances in the field. This picture is only completely tenable in the non-relativistic limit of the classical theory, however. At relativistic energies, the existence of particles, and their numbers, are fuzzy concepts which need to be given meaning by the quantum theory of fields.

1.3 Forces

In classical mechanics, forces act on particles to change their momentum. The mechanical force is defined by

$$\mathbf{F} = \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}, \tag{1.1}$$

where **p** is the momentum. In field theory, the notion of a dynamical influence is more subtle and has much in common with the interference of waves. The idea of a force is of something which acts at a point of contact and creates an impulse. This is supplanted by the notion of fields, which act at a distance and interfere with one another, and currents, which can modify the field in more subtle ways. Effective mechanical force is associated with a quantity called the *energy–momentum* tensor $\theta_{\mu\nu}$ or $T_{\mu\nu}$.

1.4 Structural elements of a dynamical system

The shift of focus, in modern physics, from particle theories to field theories means that many intuitive ideas need to be re-formulated. The aim of this book is to give a substantive meaning to the physical attributes of fields, at the classical level, so that the fully quantized theory makes physical sense. This requires example.

A detailed description of dynamical systems touches on a wide variety of themes, drawing on ideas from both historical and mathematical sources. The simplicity of field theory, as a description of nature, is easily overwhelmed by these details. It is thus fitting to introduce the key players, and mention their significance, before the clear lines of physics become obscured by the topography of a mathematical landscape. There are two kinds of dynamical system, which may be called continuous and discrete, or holonomic and non-holonomic. In this book, only systems which are parametrized by continuous, spacetime parameters are dealt with. There are three major ingredients required in the formulation of such a dynamical system.

- **Assumptions**
  A model of nature embodies a body of *assumptions* and *approximations*. The assumptions define the ultimate extent to which the theory may be considered valid. The best that physics can do is to find an idealized description of isolated phenomena under special conditions. These conditions need to be borne clearly in mind to prevent the mathematical machinery from straying from the intended path.

- **Dynamical freedom**
  The capacity for a system to change is expressed by introducing *dynamical variables*. In this case, the dynamical variables are normally fields. The number of ways in which a physical system can change is called its number of *degrees of freedom*. Such freedom describes nothing unless one sculpts out a limited form from the amorphous realm of possibility. The structure of a dynamical system is a balance between freedom and constraint.

  The variables in a dynamical system are fields, potentials and sources. There is no substantive distinction between field, potential and source, these are all simply functions of space and time; however, the words *potential* or *source* are often reserved for functions which are either static or rigidly defined by boundary conditions, whereas *field* is reserved for functions which change dynamically according to an equation of motion.

- **Constraints**
  Constraints are restrictions which determine what makes one system with n variables different from another system with n variables. The constraints of a system are both dynamical and kinematical.

  - **Equations of motion**
    These are usually the most important constraints on a system. They tell us that the dynamical variables cannot take arbitrary values; they are dynamical constraints which express limitations on the way in which dynamical variables can change.

- **Sources: external influences**

  Physical models almost always describe systems which are isolated from external influences. Outside influences are modelled by introducing *sources and sinks*. These are perturbations to a closed system of dynamical variables whose value is specified by some external boundary conditions. Sources are sometimes called generalized forces. Normally, one assumes that a source is a kind of 'immovable object' or infinite bath of energy whose value cannot be changed by the system under consideration. Sources are used to examine what happens under controlled boundary conditions. Once sources are introduced, conservation laws may be disturbed, since a source effectively opens a system to an external agent.

- **Interactions**

  Interactions are couplings which relate changes in one dynamical variable to changes in another. This usually occurs through a coupling of the equations of motion. Interaction means simply that one dynamical variable changes another. Interactions can also be thought of as internal sources, internal influences.

- **Symmetries and conservation laws**

  If a physical system possesses a symmetry, it indicates that even though one might try to affect it in a specific way, nothing significant will happen. Symmetries exert passive restrictions on the behaviour of a system, i.e. kinematical constraints. The conservation of book-keeping parameters, such as energy and momentum, is related to symmetries, so geometry and conservation are, at some level, related topics.

The Lagrangian of a dynamical theory must contain time derivatives if it is to be considered a dynamical theory. Clearly, if the rate of change of the dynamical variables with time is zero, nothing ever happens in the system, and the most one can do is to discuss steady state properties.

2
The electromagnetic field

Classical electrodynamics serves both as a point of reference and as the point of departure for the development of covariant field theories of matter and radiation. It was the observation that Maxwell's equations predict a universal speed of light *in vacuo* which led to the special theory of relativity, and this, in turn, led to the importance of *perspective* in identifying generally applicable physical laws. It was realized that the symmetries of special relativity meant that electromagnetism could be reformulated in a compact form, using a vector notation for spacetime unified into a single parameter space. The story of covariant fields therefore begins with Maxwell's four equations for the electromagnetic field in $3+1$ dimensions.

2.1 Maxwell's equations

In their familiar form, Maxwell's equations are written (in SI units)

$$\vec{\nabla} \cdot \mathbf{E} = \frac{\rho_e}{\epsilon_0} \tag{2.1a}$$

$$\vec{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.1b}$$

$$\vec{\nabla} \cdot \mathbf{B} = 0 \tag{2.1c}$$

$$c^2(\vec{\nabla} \times \mathbf{B}) = \frac{\mathbf{J}}{\epsilon_0} + \frac{\partial \mathbf{E}}{\partial t}. \tag{2.1d}$$

ρ_e is the charge density, $\mathbf{J}$ is the electric current density and $c^2 = (\epsilon_0\mu_0)^{-1}$ is the speed of light in a vacuum squared. These are valid, as they stand, in inertial frames in flat $(3+1)$ dimensional spacetimes. The study of covariant field theory begins by assuming that these equations are true, in the sense that any physical laws are 'true' – i.e. that they provide a suitably idealized description of the physics of electromagnetism. We shall not attempt to follow the path which led to their discovery, nor explore their limitations. Rather, we are interested in summarizing their form and substance, and in identifying symmetries which allow them to be expressed in an optimally simple form. In this way, we hope to learn something deeper about their meaning, and facilitate their application.

2.1.1 Potentials

This chapter may be viewed as a demonstration of how applied covariance leads to a maximally simple formulation of Maxwell's equations. A more complete understanding of electromagnetic covariance is only possible after dealing with the intricacies of chapter 9, which discusses the symmetry of spacetime. Here, the aim is to build an algorithmic understanding, in order to gain a familiarity with key concepts for later clarification.

In texts on electromagnetism, Maxwell's equations are solved for a number of problems by introducing the idea of the vector and scalar potentials. The potentials play an important role in modern electrodynamics, and are a convenient starting point for introducing covariance.

The electromagnetic potentials are introduced by making use of two theorems, which allow Maxwell's equations to be re-written in a simplified form. In a covariant formulation, one starts with these and adds the idea of a unified *spacetime*. Spacetime is the description of space and time which treats the apparently different parameters x and t in a symmetrical way. It does not claim that they are equivalent, but only that they may be treated together, since both describe different aspects of the *extent* of a system. The procedure allows us to discover a simplicity in electromagnetism which is not obvious in eqns. (2.1).

The first theorem states that the vanishing divergence of a vector implies that it may be written as the curl of some other vector quantity $\mathbf{A}$:

$$\vec{\nabla} \cdot \mathbf{v} = 0 \quad \Rightarrow \quad \mathbf{v} = \vec{\nabla} \times \mathbf{A}. \tag{2.2}$$

The second theorem asserts that the vanishing of the curl of a vector implies that it may be written as the gradient of some scalar ϕ:

$$\vec{\nabla} \times \mathbf{v} = 0 \quad \Rightarrow \quad \mathbf{v} = \vec{\nabla}\phi. \tag{2.3}$$

The deeper reason for both these theorems, which will manifest itself later, is that the curl has an *anti-symmetric* property. The theorems, as stated, are true in a homogeneous, isotropic, flat space, i.e. in a system which does not have irregularities, but they can be generalized to any kind of space. From these, one defines two *potentials*: a vector potential A_i and a scalar ϕ, which are auxiliary functions (fields) of space and time.

The physical electromagnetic field is the derivative of the potentials. From eqn. (2.1c), one defines

$$\mathbf{B} = \vec{\nabla} \times \mathbf{A}. \tag{2.4}$$

This form completely solves that equation. One equation has now been automatically and completely solved by re-parametrizing the problem in terms of a new variable. Eqn. (2.1c) tells us now that

$$\vec{\nabla} \times \mathbf{E} = -\frac{\partial}{\partial t}(\vec{\nabla} \times \mathbf{A})$$

$$\vec{\nabla} \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0. \tag{2.5}$$

Consequently, according to the second theorem, one can write

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\vec{\nabla}\phi, \tag{2.6}$$

giving

$$\mathbf{E} = -\vec{\nabla}\phi - \frac{\partial \mathbf{A}}{\partial t}. \tag{2.7}$$

The minus sign on the right hand side of eqn. (2.6) is the convention which is used to make attractive forces positive and repulsive forces negative.

Introducing potentials in this way is not a necessity: many problems in electromagnetism can be treated by solving eqns. (2.1) directly, but the introduction often leads to significant simplifications when it is easier to solve for the potentials than it is to solve for the fields.

The potentials themselves are a mixed blessing: on the one hand, the re-parametrization leads to a number of helpful insights about Maxwell's equations. In particular, it reveals *symmetries*, such as the gauge symmetry, which we shall explore in detail later. It also allows us to write the matter–radiation interaction in a *local* form which would otherwise be impossible. The price one pays for these benefits is the extra conceptual layers associated with the potential and its gauge invariance. This confuses several issues and forces us to deal with constraints, or conditions, which uniquely define the potentials.

2.1.2 Gauge invariance

Gauge invariance is a symmetry which expresses the freedom to re-define the potentials arbitrarily without changing their physical significance. In view of the theorems above, the fields **E** and **B** are invariant under the re-definitions

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \vec{\nabla}s$$

$$\phi \to \phi' = \phi - \frac{\partial s}{\partial t}. \tag{2.8}$$

These re-definitions are called *gauge transformations*, and $s(x)$ is an arbitrary scalar function. The transformation means that, when the potentials are used