3
Field parameters

The parameters which measure change in dynamical systems have a unique importance: they describe both the layout and the development of a system. Space (position) and time are the most familiar parameters, but there are other possibilities, such as Fourier modes.

In the previous chapter, it was seen how the unification of spatial and temporal parameters, in electromagnetism, led to a tidier and deeper form of the Maxwell equations. It also made the equations easier to transform into other relativistic frames. In the covariant approach to physics one is concerned with what does and does not change, when shifting from one perspective to another, i.e. with the properties of a system which are dependent and independent of the circumstances of observation. In a continuous, holonomic system, this is summarized by two independent concepts: parameter spaces and coordinates.

- **Parameter space** (manifold). This represents the stage for physical reality. A parameter space has coordinate-independent properties such as topology and curvature.

- **Coordinates**. These are arbitrary labels used to mark out a reference scheme, or measurement scheme, in parameter space. There is no unique way to map out a parameter space, e.g. Cartesian or polar coordinates. If there is a special symmetry, calculations are often made easier by choosing coordinates which match this symmetry.

Coordinates are labels which mark a scale on a parameter space. They measure a distance in a particular direction from an arbitrary origin. Clearly, there is nothing fundamental about coordinates: by changing the arbitrary origin, or orientation of measurement, all coordinate labels are changed, but the underlying reality is still the same. They may be based on flat Cartesian (x, y, z) or polar (r, θ, ϕ) conventions; they can be marked on flat sheets or curved shells.

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

Underneath the details of an arbitrary system of measurement is a physical system which owes nothing to those details.

The invariant properties or *symmetries* of parameter spaces have many implicit consequences for physical systems; not all are immediately intuitive. For this reason, it is useful to study these invariant properties in depth, to see how they dictate the possibilities of behaviour (see chapter 9). For now it is sufficient to define a notation for coordinates on the most important parameter spaces.

This chapter summarizes the formulation of $(n + 1)$ dimensional vectors in Minkowski spacetime and in its complementary space of wavevectors k, usually called *momentum space* or *reciprocal lattice space*.

3.1 Choice of parametrization

The dynamical variables, in field theory, are the fields themselves. They are functions of the parameters which map out the background space or spacetime; e.g.

$$\psi(t),\ \phi(t, \mathbf{x}),\ \chi(t, r, \theta, \phi). \tag{3.1}$$

Field variables are normally written as functions of spacetime positions, but other decompositions of the field are also useful. Another ubiquitous choice is to use a complementary set of variables based upon a decomposition of the field into a set of basis functions, a so-called *spectral decomposition*. Given a complete set of functions $\psi_i(x)$, one can always write an arbitrary field as a linear super-position:

$$\phi(x) = \sum_i c_i\, \psi_i(x). \tag{3.2}$$

Since the functions are fixed and known, a knowledge of the coefficients c_i in this decomposition is equivalent to a knowledge of $\phi(x)$, i.e. as a function of x. However, the function may also be written in a different parametrization:

$$\phi(c_1, c_2, c_3 \ldots). \tag{3.3}$$

This is a shorthand for the decomposition above, just as $\phi(x)$ is a shorthand for a polynomial or series in x. Usually, an infinite number of such coefficients is needed to prescribe a complete decomposition of the field, as, for instance, in the Fourier expansion of a function, described below.

Spacetime is an obvious parameter space for a field theory since it comprises the world around us and it includes laboratories where experiments take place, but other basis functions sometimes reveal simpler descriptions. One important example is the complementary Fourier transform of spacetime. The Fourier

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

transform is important in situations where one suspects a translationally invariant, homogeneous system. The Fourier transform of a function of x is defined to be a new function of the wavenumber k (and the inverse transform) by the relations:

$$f(x) = \int \frac{\mathrm{d}k}{2\pi}\, \mathrm{e}^{\mathrm{i}kx}\, f(k)$$
$$f(k) = \int \mathrm{d}x\, \mathrm{e}^{\mathrm{i}kx}\, f(x). \tag{3.4}$$

k is a continuous label on a continuous set of functions $\exp(\mathrm{i}kx)$, not a discrete set of c_i, for integer i. In solid state physics, the space parametrized by k is called the reciprocal lattice space. Fourier transform variables are useful for many purposes, such as revealing hidden periodicities in a function, since the expansion is based on periodic functions. The Fourier transform is also a useful calculational aid.

Spacetime (configuration space) and the Fourier transform are two complementary ways of describing the basic evolution of most systems. These two viewpoints have advantages and disadvantages. For example, imagine a two-state system whose behaviour in time can be drawn as a square wave. To represent a square wave in Fourier space, one requires either an infinite number of Fourier waves of different frequencies, or merely two positions over time. In that case, it would be cumbersome to use a Fourier representation of the time evolution.

3.2 Configuration space

The four-dimensional vectors used to re-write electromagnetism are easily generalized to $(n+1)$ spacetime dimensions, for any positive n. They place time and space on an *almost* equal footing. In spite of the notational convenience of unified spacetime, some caution is required in interpreting the step. Time is *not* the same as space: formally, it distinguishes itself by a sign in the metric tensor; physically, it plays a special role in determining the dynamics of a system.

3.2.1 Flat and curved space

Physical systems in constrained geometries, such as on curved surfaces, or within containers, are best described using curvilinear coordinates. Experimental apparatus is often spherical or toroidal; shapes with a simple symmetry are commonly used when generating electromagnetic fields; rectangular fields with sharp corners are less common, since these require much higher energy to sustain.

Studies of what happens within the volumes of containers, and what happens on their surface boundaries, are important in many situations [121]. When

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

generalizing, to study systems in $(n + 1)$ dimensions, the idea of surfaces and volumes also has to be generalized. The distinction becomes mainly one of convenience: $(n + 1)$ dimensional curved surfaces are curved spacetimes. The fact that they enclose a volume or partition a space which is $(n+2)$ dimensional is not always germane to the discussion at hand. This is particularly true in cosmology.

It is important to distinguish between curvilinear coordinates in flat space and coordinate labellings of curved space. An example of the former is the use of polar (r, θ) coordinates to map out a plane. The plane is flat, but the coordinates span the space in a set of curved rings. An example of the latter is (θ, ϕ) coordinates (at fixed r), mapping out the surface of a sphere. Over very short distances, (θ, ϕ) can be likened to a tiny planar patch with Cartesian coordinates (x, y).

Einstein's contribution to the theory of gravity was to show that the laws of gravitation could be considered as an intrinsic curvature of a $(3+1)$ dimensional spacetime. Einstein used the idea of covariance to argue that one could view gravity in one of two equivalent ways: as forced motion in a flat spacetime, or as free-fall in a curved spacetime. Using coordinates and metric tensors, gravitation could itself be described as a field theory, in which the field $g_{\mu\nu}(x)$ was the shape of spacetime itself.

Gravitational effects may be built into a covariant formalism to ensure that every expression is general enough to be cast into an arbitrary scheme of coordinates. If one allows for general coordinates (i.e. general covariance), one does not assume that all coordinates are orthogonal Cartesian systems, and gravity and curvature are not excluded from the discussion.

Spacetime curvature will not be treated in detail here, since this topic is widely discussed in books on relativity. However, we take the issue of curvature 'under advisement' and construct a formalism for dealing with arbitrary coordinates, assured that the results will transform correctly even in a curved environment.

3.2.2 Vector equations

Vector methods express spatial relationships, which remain true regardless of the system of coordinates used to write them down. They thus play a central role in covariant formulation. For example, the simple vector equation

$$\mathbf{A} \cdot \mathbf{B} = 0 \tag{3.5}$$

expresses the fact that two vectors $\mathbf{A}$ and $\mathbf{B}$ are orthogonal. It says nothing about the orientation of the vectors relative to a coordinate system, nor their position relative to an origin; rather, it expresses a relationship of more intrinsic value between the vectors: their relative orientation. Vector equations and covariance are natural partners.

Vector equations are form-invariant under changes of coordinates, but the details of their components do change. For instance, in the above equation, if one fixes a coordinate system, then the components of the two vectors take on definite values. If one then rotates or translates the coordinates, the values of the components change, but the equation itself remains true.

3.2.3 Coordinate bases

A coordinate basis is a set of $(n + 1)$ linearly independent reference vectors $\mathbf{e}_\mu$, used to provide a concise description of any vector within a vector space. They are 'standard arrows'; without them, every direction would need to have a different name.[1]

In index notation, the components of a vector $\mathbf{a}$ are written, relative to a basis or set of axes $\mathbf{e}_i$, as $\{a^i\}$, i.e.

$$\mathbf{a} = \sum_\mu a^\mu \, \mathbf{e}_\mu \equiv a^\mu \, \mathbf{e}_\mu . \tag{3.6}$$

Note that, as usual, there is an implied summation convention over repeated indices throughout this book. The subscript μ runs over the number of dimensions of the space.

Linearity is a central concept in vector descriptions. One does not require what happens within the space to be linear, but the basis vectors must be locally linear in order for the vector description to be single-valued. Consider, then, the set of all linear scalar functions of vectors. Linearity implies that a linear combination of arguments leads to a linear combination of the functions:

$$\omega(c^\mu \mathbf{e}_\mu) = c^\mu \omega(\mathbf{e}_\mu). \tag{3.7}$$

Also, the linear combination of different functions results in new linear functions:

$$\omega'(\mathbf{v}) = \sum_\mu c_\mu \omega^\mu(\mathbf{v}). \tag{3.8}$$

The space of these functions is therefore also a vector space V^*, called the dual space. It has the same dimension as the vector space (also called the tangent space). The duality refers to the fact that one may consider the 1-forms to be linear functions of the basis vectors, or vice versa, i.e.

$$\omega(\mathbf{v}) = \mathbf{v}(\omega). \tag{3.9}$$

[1] In terms of information theory, the vector basis provides a systematic $(n+1)$-tuple of numbers, which in turn provides an optimally compressed coding of directional information in the vector space. Without such a system, we would be stuck with names like north, south, east, west, north-north-west, north-north-north-west etc. for each new direction.

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

Vector components v^i are written

$$\mathbf{v} = v^\mu \mathbf{e}_\mu, \tag{3.10}$$

and dual vector (1-form) components are written

$$\mathbf{v} = v_\mu \omega^\mu. \tag{3.11}$$

The scalar product is

$$\begin{aligned} \mathbf{v} \cdot \mathbf{v} \;=\; \mathbf{v}^* \mathbf{v} &= (v_\mu \omega^\mu)(v^\nu \mathbf{e}_\nu) \\ &= v_\mu v^\nu \; (\omega^\mu \mathbf{e}_\nu) \\ &= v_\mu v^\nu \; \delta_\mu^{\ \nu} \\ &= v_\mu v^\mu, \end{aligned} \tag{3.12}$$

where

$$(\omega^\mu \mathbf{e}_\nu) = \delta^\mu_{\ \nu}. \tag{3.13}$$

The metric tensor $g_{\mu\nu}$ maps between these equivalent descriptions:

$$\begin{aligned} v_\mu &= g_{\mu\nu} v^\nu \\ v^\mu &= g^{\mu\nu} v_\nu, \end{aligned} \tag{3.14}$$

and

$$\mathbf{e}_\mu \cdot \mathbf{e}_\nu = g_{\mu\nu} \tag{3.15a}$$

$$\omega^\mu \cdot \omega^\nu = g^{\mu\nu}. \tag{3.15b}$$

When acting on scalar functions, the basis vectors $\mathbf{e}_\mu \to \partial_\mu$ are tangential to the vector space; the 1-forms $\omega^\mu \to \mathrm{d}x^\mu$ lie along it.

In general, under an infinitesimal shift of the coordinate basis by an amount $\mathrm{d}x^\mu$, the basis changes by an amount

$$\mathrm{d}\mathbf{e}_\mu = \Gamma_{\mu\nu}^{\ \ \lambda} \mathbf{e}_\lambda \, \mathrm{d}x^\nu. \tag{3.16}$$

The symbol $\Gamma_{\mu\nu}^{\ \ \lambda}$ is called the affine connection, or Christoffel symbol. From this, one determines that

$$\partial_\nu \mathbf{e}_\mu = \Gamma_{\mu\nu}^{\ \ \lambda} \mathbf{e}_\lambda, \tag{3.17}$$

and by differentiating eqn. (3.13), one finds that

$$\partial_\nu \omega^\lambda = -\Gamma_{\nu\mu}^{\ \ \lambda} \omega_\mu. \tag{3.18}$$

The connection can be expressed in terms of the metric, by differentiating eqn. (3.15a):

$$\begin{aligned} \partial_\lambda g_{\mu\nu} &= \partial_\lambda \mathbf{e}_\mu \cdot \mathbf{e}_\nu + \mathbf{e}_\mu \cdot \partial_\lambda \mathbf{e}_\nu \\ &= \Gamma_{\mu\lambda}^{\ \ \lambda} g_{\rho\nu} + g_{\rho\mu} \Gamma_{\nu\lambda}^{\ \ \lambda}. \end{aligned} \tag{3.19}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

By permuting indices in this equation, one may show that

$$\Gamma_{\lambda\mu}{}^{\sigma} = \frac{1}{2} g^{\nu\sigma} \left\{ \partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\nu g_{\mu\lambda} \right\}. \tag{3.20}$$

The connection is thus related to cases where the metric tensor is not constant. This occurs in various contexts, such when using curvilinear coordinates, and when fields undergo conformal transformations, such as in the case of gauge transformations.

3.2.4 Example: Euclidean space

In n-dimensional Euclidean space, the spatial indices i of a vector's components run from 1 to n except where otherwise stated. The length of a vector interval **ds** is an invariant quantity, which is defined by the inner product. This may be written

$$\mathbf{ds} \cdot \mathbf{ds} = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \tag{3.21}$$

in a Cartesian basis. In the index notation (for $n = 3$) this may be written,

$$\mathbf{ds} \cdot \mathbf{ds} = \mathrm{d}x^i \mathrm{d}x_i. \tag{3.22}$$

Repeated indices are summed over, unless otherwise stated. We distinguish, in general, between vector components with raised indices (called *contravariant* components) and those with lower indices (called, confusingly, *covariant* components,[2] and 'normal' components, which we shall almost never use. In a Cartesian basis $(x, y, z \ldots)$ there is no difference between these components. In other coordinate systems, such as polar coordinates however, they are different.

Results which are independent of coordinate basis always involve a sum over one raised index and one lower index. The length of the vector interval above is an example. We can convert an up index into a down index using a matrix (actually a tensor) called the metric tensor g_{ij},

$$a_i = g_{ij} a^j. \tag{3.23}$$

The inverse of the metric g_{ij} is written g^{ij} (with indices raised), and it serves to convert a lower index into an upper one:

$$a^i = g^{ij} a_j. \tag{3.24}$$

The metric and its inverse satisfy the relation,

$$g_{ij} g^{jk} = g_i{}^k = \delta_i{}^k. \tag{3.25}$$

[2] There is no connection between this designation and the usual meaning of covariant.

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

In Cartesian components, the components of the metric are trivial. It is simply the identity matrix, or Kronecker delta:

$$\text{(Cartesian)}: \quad g_{ij} = g^{ij} = \delta_{ij}. \tag{3.26}$$

To illustrate the difference between covariant, contravariant and normal components, consider two-dimensional polar coordinates as an example. The vector interval, or line element, is now written

$$\mathbf{ds} \cdot \mathbf{ds} = \mathrm{d}r^2 + r^2 \mathrm{d}\theta^2. \tag{3.27}$$

The normal components of the vector **ds** have the dimensions of length in this case, and are written

$$(\mathrm{d}r, r\,\mathrm{d}\theta). \tag{3.28}$$

The contravariant components are simply the coordinate intervals,

$$\mathrm{d}s^i = (\mathrm{d}r, \mathrm{d}\theta), \tag{3.29}$$

and the covariant components are

$$\mathrm{d}s_i = (\mathrm{d}r, r^2\mathrm{d}\theta). \tag{3.30}$$

The metric tensor is then defined by

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \tag{3.31}$$

and the inverse tensor is simply

$$g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}. \tag{3.32}$$

The covariant and contravariant components are used almost exclusively in the theory of special relativity.

Having introduced the metric tensor, we may define the scalar product of any two vectors **a** and **b** by

$$\mathbf{a} \cdot \mathbf{b} = a^i b_i = a^i g_{ij} b^j. \tag{3.33}$$

The definition of the vector product and the curl are special to three space dimensions. We define the completely anti-symmetric tensor in three dimensions by

$$\epsilon^{ijk} = \begin{cases} +1 & ijk = 123 \text{ and even permutations} \\ -1 & ijk = 321 \text{ and other odd permutations} \\ 0 & \text{otherwise.} \end{cases} \tag{3.34}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

This is also referred to as the three-dimensional Levi-Cevita tensor in some texts. Since its value depends on permutations of 123, and its indices run only over these values, it can only be used to generate products in three dimensions. There are generalizations of this quantity for other numbers of dimensions, but the generalizations must always have the same number of indices as spatial dimensions, thus this object is unique in three dimensions. More properties of anti-symmetric tensors are described below.

In terms of this tensor, we may write the ith covariant component of the three-dimensional vector cross-product as

$$(\mathbf{b} \times \mathbf{c})_i = \epsilon_{ijk} b^j c^k. \tag{3.35}$$

Contracting with a scalar product gives the volume of a parallelepiped spanned by vectors **a**, **b** and **c**,

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \epsilon_{ijk} a^i b^j c^k, \tag{3.36}$$

which is basis-independent.

3.2.5 Example: Minkowski spacetime

The generalization of Euclidean space to relativistically motivated spacetime is called Minkowski spacetime. Close to the speed of light, the lengths of n-dimensional spatial vectors are not invariant under boosts (changes of speed), due to the Lorentz length contraction. From classical electromagnetism, one finds that the speed of light in a vacuum must be constant for all observers:

$$c^2 = \frac{1}{\epsilon_0 \mu_0}, \tag{3.37}$$

and one deduces from this that a new quantity is invariant; we refer to this as the invariant line element

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = -c^2\,\mathrm{d}\tau^2, \tag{3.38}$$

where $\mathrm{d}\tau$ is referred to as the proper time. By comparing the middle and rightmost terms in this equation, it may be seen that the proper time is the time coordinate in the rest frame of a system, since there is no change in the position variables. The negative sign singles out the time contribution as special. The nomenclature 'timelike separation' is used for intervals in which $\mathrm{d}s^2 < 0$, 'spacelike separation' is used for $\mathrm{d}s^2 > 0$, and 'null' is used for $\mathrm{d}s^2 = 0$.

In terms of $(n+1)$ dimensional vectors, one writes:

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \mathrm{d}x_\mu = \mathrm{d}x^\mu g_{\mu\nu} \mathrm{d}x^\nu \tag{3.39}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

where $\mu, \nu = 0, 1, 2, \ldots, n$ In a Cartesian basis, the contravariant and covariant components of the spacetime interval are defined, respectively, by

$$\begin{aligned} \mathrm{d}x^{\mu} &= (\, ct, x, y, z, \ldots) \\ \mathrm{d}x_{\mu} &= (-ct, x, y, z, \ldots), \end{aligned} \tag{3.40}$$

and the metric tensor in this Cartesian basis, or locally inertial frame (LIF), is the constant tensor

$$\eta_{\mu\nu} \equiv g_{\mu\nu}\Big|_{\mathrm{LIF}} = \begin{pmatrix} -1 & 0 & 0\cdots & 0 \\ 0 & 1 & 0\cdots & 0 \\ 0 & 0 & 1\cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0\cdots & 1 \end{pmatrix}. \tag{3.41}$$

This is a special case of a metric in a general frame $g_{\mu\nu}$.

This placement of signs in the metric is arbitrary, and two other conventions are found in the literature: the opposite sign for the metric, with corresponding movement of the minus sign from the time to the space parts in the covariant and contravariant components; and a Euclidean formulation, in which the metric is entirely positive (positive definite), and the time components of the components are symmetrically ict. This last form, called a Euclidean formulation (or Riemannian in curved spacetime), has several uses, and thus we adopt conventions in this text in which it is trivial to convert to the Euclidean form and back.

Contravariant vectors describe regular parametrizations of the coordinates. In order to define a frame-invariant derivative, we need to define partial derivatives by the requirement that the partial derivative of x^1 with respect to x_1 be unity:

$$\frac{\partial}{\partial x^1} x^1 = \partial_1 x^1 = 1. \tag{3.42}$$

Notice that 'dividing by' an upper index makes it into an object with an effectively lower index. More generally, we require:

$$\frac{\partial}{\partial x^{\mu}} x^{\nu} = \partial_{\mu} x^{\nu} = \delta_{\mu}{}^{\nu}. \tag{3.43}$$

From this, one sees that the Cartesian components of the derivative must be

$$\begin{aligned} \partial_{\mu} &= \left(\frac{1}{c}\partial_t, \partial_x, \partial_y, \partial_z, \ldots \right) \\ \partial^{\mu} &= \left(-\frac{1}{c}\partial_t, \partial_x, \partial_y, \partial_z, \ldots \right). \end{aligned} \tag{3.44}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

Velocity is a relative concept, by definition. It is intimately associated with a choice of Lorentz frame. The relative velocity is defined as the time derivative of the position

$$\beta^\mu = \frac{1}{c}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \frac{\mathrm{d}x^\mu}{\mathrm{d}x^0}. \tag{3.45}$$

Unfortunately, because both x^μ and t are frame-dependent, this quantity does not transform like a vector. To obtain a vector, we choose to look at

$$U^\mu = \frac{1}{c}\frac{x^\mu}{\mathrm{d}\tau}. \tag{3.46}$$

The components of the relative velocity are as follows:

$$\beta^\mu = (\beta^0, \beta^i) = (1, v^i/c). \tag{3.47}$$

The relationship to the velocity vector is given by

$$U^\mu = \gamma c \beta^\mu. \tag{3.48}$$

Hence,

$$U^\mu U_\mu = -c^2. \tag{3.49}$$

3.3 Momentum space and waves

The reciprocal wavevector space of k_μ plays a complementary role to that of spacetime. It measures changes in waves when one is not interested in spacetime locations. Pure harmonic (sinusoidal) waves are spread over an infinite distance. They have no beginning or end, only a definite wavelength.

In the quantum theory, energy and momentum are determined by the operators

$$E \to \mathrm{i}\hbar\partial_t, \quad p_i \to -\mathrm{i}\hbar\partial_i, \tag{3.50}$$

which have pure values when acting on plane wave states

$$\psi \sim \exp \mathrm{i}(k_i x^i - \omega t). \tag{3.51}$$

In $(n+1)$ dimensional notation, the wavevector becomes:

$$k_\mu = \left(-\frac{\omega}{c}, k_i\right), \tag{3.52}$$

so that plane waves take the simple form

$$\psi \sim \exp(\mathrm{i}k_\mu x^\mu). \tag{3.53}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

The energy and momentum are therefore given by the time and space eigenvalues of the operator

$$p_\mu = -\mathrm{i}\hbar\partial_\mu, \tag{3.54}$$

respectively, as they act upon a plane wave. This leads to the definition of an $(n+1)$ dimensional energy–momentum vector,

$$p_\mu = \hbar k_\mu = \left(-\frac{E}{c}, p_i\right). \tag{3.55}$$

The identification $p_\mu = \hbar k_\mu$ is the de Broglie relation for matter waves. This is one of the most central and important relations in the definition of the quantum theory of matter.

In discussing wavelike excitations, it is useful to resolve the components of vectors along the direction of motion of the wave (longitudinal) and perpendicular (transverse) to the direction of motion. A longitudinal vector is one proportional to a vector in the direction of motion of a wave k^μ. A transverse vector is orthogonal to this vector. The longitudinal and transverse components of a vector are defined by

$$\begin{aligned} V_{\mathrm{L}}^\mu &\equiv \frac{k^\mu k_\nu}{k^2}\, V^\nu \\ V_{\mathrm{T}}^\mu &\equiv \left(g_{\mu\nu} - \frac{k^\mu k_\nu}{k^2}\right) V^\nu. \end{aligned} \tag{3.56}$$

It is straightforward to verify that the two projection operators

$$\begin{aligned} {P_{\mathrm{L}}}^\mu_{\ \nu} &= \frac{k^\mu k_\nu}{k^2} \\ {P_{\mathrm{T}}}^\mu_{\ \nu} &= \left(g_{\mu\nu} - \frac{k^\mu k_\nu}{k^2}\right) \end{aligned} \tag{3.57}$$

are orthogonal to one another:

$$(P_{\mathrm{L}})^\mu_{\ \nu}(P_{\mathrm{T}})^\nu_{\ \lambda} = 0. \tag{3.58}$$

3.4 Tensor transformations

Vector equations remain true in general coordinate frames because the components of a vector transform according to specific rules under a coordinate transformation U:

$$\mathbf{v}' = U\,\mathbf{v}, \tag{3.59}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

or

$$v'^i = U^i{}_j \, v^j, \tag{3.60}$$

where the components of the matrix U are fixed by the requirement that the equations remain true in general coordinate systems. This is a valuable property, and we should be interested in generalizations of this idea which might be useful in physics.

Tensors are objects with any number of indices, which have the same basic transformation properties as vectors. The number of indices on a tensor is its rank. Each free index in a tensor equation requires a transformation matrix under changes of coordinates; free indices represent the components in a specific coordinate basis, and each summed index is invariant since scalar products are independent of basis.

Under a change of coordinates, $x \to x'$, a scalar (rank 0-tensor) transforms simply as

$$\phi(x) \to \phi(x'). \tag{3.61}$$

For a vector (rank 1-tensor), such a simple rule does make sense. If one rotates a coordinate system, for instance, then all the components of a vector must change, since it points in a new direction with respect to the coordinate axes. Thus, a vector's components must transform separately, but as linear combinations of the old components. The rule for a vector with raised index is:

$$V^\mu(x') = \frac{\partial x'^\mu}{\partial x^\nu} \, V^\nu(x) = (\partial_\nu x'^\mu) \, V^\nu(x). \tag{3.62}$$

For a vector with lowered index, it is the converse:

$$V_\mu(x') = \frac{\partial x^\nu}{\partial x'^\mu} \, V_\nu(x) = (\partial'_\mu x^\nu) \, V_\nu(x). \tag{3.63}$$

Here we have used two notations for the derivatives: the longhand notation first for clarity and the shorthand form which is more compact and is used throughout this book.

The metric tensor is a tensor of rank 2. Using the property of the metric in raising and lowering indices, one can also deduce its transformation rule under the change of coordinates from x to x'. Starting with

$$V^\mu(x') = g^{\mu\nu}(x') V_\nu(x'), \tag{3.64}$$

and expressing it in the x coordinate system, using the transformation above, one obtains:

$$(\partial_\nu x'^\mu) V^\nu(x) = g^{\mu\sigma}(x') (\partial'_\sigma x^\rho) V_\rho(x). \tag{3.65}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

However, it is also known that, in the unprimed coordinates,

$$V^{\nu}(x) = g^{\nu\sigma}(x)V_{\sigma}(x). \tag{3.66}$$

Comparing eqns. (3.65) and (3.66), it is possible to deduce the transformation rule for the inverse metric $g^{\mu\nu}$. To do this, one rearranges eqn. (3.65) by multiplying by $(\partial'_{\mu}x^{\tau})$ and using the chain-rule:

$$(\partial_{\nu}x'^{\mu})(\partial'_{\mu}x^{\tau}) = \delta_{\nu}{}^{\tau}. \tag{3.67}$$

Being careful to re-label duplicate indices, this gives

$$\delta\nu\ \tau\ V^{\nu}(x) = g^{\mu\sigma}(x')(\partial'_{\mu}x^{\tau})(\partial'_{\sigma}x^{\rho})\ V_{\rho}(x), \tag{3.68}$$

which is

$$V^{\tau}(x) = g^{\mu\rho}(x')(\partial'_{\mu}x^{\tau})(\partial'_{\rho}x^{\sigma})V_{\sigma}(x). \tag{3.69}$$

Comparing this with eqn. (3.66), one finds that

$$g^{\rho\mu}(x')(\partial'_{\mu}x^{\tau})(\partial'_{\rho}x^{\sigma}) = g^{\tau\sigma}(x), \tag{3.70}$$

or, equivalently, after re-labelling and re-arranging once more,

$$g^{\mu\nu}(x') = (\partial_{\rho}x'^{\mu})(\partial_{\sigma}x'^{\nu})g^{\rho\sigma}(x). \tag{3.71}$$

One sees that this follows the same pattern as the vector transformation with raised indices. The difference is that there is now a partial derivative matrix $(\partial_{\sigma}x'^{\nu})$ for each index. In fact, this is a general feature of tensors. Each raised index transforms with a factor like $(\partial_{\sigma}x'^{\nu})$ and each lowered index transforms with a factor like $\partial'_{\mu}x^{\nu}$. For instance,

$$T^{\mu\nu}{}_{\rho\sigma}(x') = (\partial_{\alpha}x'^{\mu})(\partial_{\beta}x'^{\nu})(\partial'_{\rho}x^{\gamma})(\partial'_{\sigma}x^{\delta})T^{\alpha\beta}{}_{\gamma\delta}. \tag{3.72}$$

3.5 Properties

The following properties of tensors are instructive and useful.

(1) Any matrix T may be written as a sum of a symmetric part $\overline{T}_{ij} = \frac{1}{2}(T_{ij} + T_{ji})$ and an anti-symmetric part $\tilde{T}_{ij} = \frac{1}{2}(T_{ij} - T_{ji})$. Thus one may write any 2×2 matrix in the form

$$T = \begin{pmatrix} \overline{T}_{11} & \overline{T}_{12} + \tilde{T}_{12} \\ \overline{T}_{12} - \tilde{T}_{12} & \overline{T}_{22} \end{pmatrix} \tag{3.73}$$

(2) It may be shown that the trace of the product of a symmetric matrix with an anti-symmetric matrix is zero, i.e. $\overline{S}^{ij}\tilde{T}_{ij} = 0$.

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

(3) By considering similarity transformations of the form $T \to \Lambda^{-1} T \Lambda$, one may show that the trace of any matrix is an invariant, equal to the sum of its eigenvalues.

(4) By definition, a rank 2-tensor T transforms by the following matrix multiplication rule:

$$T \to \Lambda^{\mathrm{T}} T \Lambda, \tag{3.74}$$

for some transformation matrix Λ. Consider a general 2×2 tensor

$$T = \begin{pmatrix} \frac{1}{2}t + \Delta T_{11} & \overline{T}_{12} + \tilde{T}_{12} \\ \overline{T}_{12} + \tilde{T}_{12} & \frac{1}{2}t + \Delta T_{22} \end{pmatrix},$$

where t is the trace $t = (\overline{T}_{11} + \overline{T}_{22})$, and consider the effect of the following matrices on T:

$$\begin{aligned} \Lambda_0 &= \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \\ \Lambda_1 &= \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix} \\ \Lambda_2 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \Lambda_3 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \end{aligned} \tag{3.75}$$

For each of these matrices, compute:

(a) $\Lambda^{\mathrm{T}} \Lambda$,

(b) $\Lambda^{\mathrm{T}} T \Lambda$.

It may be shown that, used as a transformation on T:

(a) the anti-symmetric matrix Λ_1 leaves anti-symmetric terms invariant and preserves the trace of T;

(b) the off-diagonal symmetric matrix Λ_2 leaves the off-diagonal symmetric terms invariant and preserves the trace of T;

(c) the symmetrical, traceless matrix Λ_3, preserves only the trace of T.

It may thus be concluded that a tensor T in n dimensions has three separately invariant parts and may be written in the form

$$T_{ij} = \frac{1}{n} T^k_{\ k} \delta_{ij} + \overline{T}_{ij} + \left(\tilde{T}_{ij} - \frac{1}{n} T^k_{\ k} \delta_{ij} \right). \tag{3.76}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

3.6 Euclidean and Riemannian spacetime

Minkowski spacetime has an indefinite metric tensor signature. In Euclidean and Riemannian spacetime, the metric signature is definite (usually positive definite). General curved spaces with a definite signature are referred to as Riemannian manifolds. Multiplying the time components of vectors and tensors by the square-root of minus one (i) allows one to pass from Minkowski spacetime to Euclidean spacetime and back again. This procedure is known as *Wick rotation* and is encountered in several contexts in quantum theory. For instance, it serves a regulatory role: integrals involving the Lorentzian form $(k^2 + m^2)^{-1}$ are conveniently evaluated in Euclidean space, where $k^2 + m^2$ has no zeros. Also, there is a convenient relationship between equilibrium thermodynamics and quantum field theory in Euclidean space.

We shall use subscripts and superscripts 'E' to indicate quantities in Euclidean space; 'M' denotes Minkowski space, for this section only. Note that the transformation affects only the time or zeroth components of tensors; the space parts are unchanged.

The transformation required to convert from Minkowski spacetime (with its indefinite metric) to Euclidean spacetime (with its definite metric) is motivated by the appearance of plane waves in the Fourier decomposition of field variables. Integrals over plane waves of the form $\exp\ \mathrm{i}(\mathbf{k}\cdot\mathbf{x} - \omega t)$ have no definite convergence properties, since the complex exponential simply oscillates for all values of $\mathbf{k}$ and ω. However, if one adds a small imaginary part to time $t \to t - \mathrm{i}\tau$, then we turn the oscillatory behaviour into exponential decay:

$$\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)} \to \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\mathbf{x}-\omega t)}\mathrm{e}^{-\omega\tau}. \tag{3.77}$$

The requirement of decay rather than growth chooses the sign for the Wick rotation. An equivalent motivation is to examine the Lorentzian form:

$$\frac{1}{k^2+m^2} = \frac{1}{-k_0^2+\mathbf{k}^2+m^2} = \frac{1}{(-k_0+\sqrt{\mathbf{k}^2+m^2})(k_0+\sqrt{\mathbf{k}^2+m^2})}. \tag{3.78}$$

This is singular and has poles on the real k_0 axis at $k_0 = \pm\sqrt{\mathbf{k}^2+m^2}$. This makes the integral of k_0 non-analytical, and a prescription must be specified for integrating around the poles. The problem can be resolved by adding a small (infinitesimal) imaginary part to the momenta:

$$\frac{1}{k^2+m^2-\mathrm{i}\epsilon} = \frac{1}{(-k_0-\mathrm{i}\epsilon+\sqrt{\mathbf{k}^2+m^2})(k_0-\mathrm{i}\epsilon+\sqrt{\mathbf{k}^2+m^2})}. \tag{3.79}$$

This effectively shifts the poles from the real axis to above the axis for negative k_0 and below the axis for positive k_0. Since it is possible to rotate the contour

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press

90 degrees onto the imaginary axis without having to pass through any poles, by defining (see section 6.1.1)

$$k_0^{\mathrm{E}} = \mathrm{i}k_0, \tag{3.80}$$

this once again chooses the sign of the rotation. The contour is rotated clockwise by 90 degrees, the integrand is positive definite and no poles are encountered in an integral over κ_0:

$$\frac{1}{-k_0^2 + \mathbf{k}^2 + m^2 - \mathrm{i}\epsilon} \to \frac{1}{k_{0\mathrm{E}}^2 + \mathbf{k}^2 + m^2}. \tag{3.81}$$

All variables in a field theory must be rotated consistently:

$$x_{\mathrm{E}}^0 = -\mathrm{i}x^0 \tag{3.82}$$
$$x_0^{\mathrm{E}} = \mathrm{i}x_0 \tag{3.83}$$
$$k_0^{\mathrm{E}} = \mathrm{i}k_0 = -\mathrm{i}\omega/c. \tag{3.84}$$

The inner product

$$k_\mu x^\mu = \mathbf{k}\cdot\mathbf{x} + k_0 x^0 \to \mathbf{k}\cdot\mathbf{x} + \kappa_0 x^0 \tag{3.85}$$

is consistent with

$$\partial_0 x^0 = \partial_0^{\mathrm{E}}\, x_{\mathrm{E}}^0 = 1 \tag{3.86}$$

where

$$\partial_0^{\mathrm{E}} = \mathrm{i}\partial_0, \tag{3.87}$$

since $\partial_0^{\mathrm{E}} \to \mathrm{i}\kappa_0$. Since the Wick transformation affects derivatives and vectors, it also affects Maxwell's equations. From

$$\partial^\nu F_{\mu\nu} = \mu_0 J_\mu, \tag{3.88}$$

we deduce that

$$J_0^{\mathrm{E}} = \mathrm{i}J_0 \tag{3.89}$$
$$A_0^{\mathrm{E}} = \mathrm{i}A_0, \tag{3.90}$$

which are necessary in view of the homogeneous form of the field strength:

$$-\mathrm{i}F_{0i}^{\mathrm{E}} = \partial_0 A_i - \partial_i A_0 = F_{0i}. \tag{3.91}$$

Notice that, in $(3+1)$ dimensions, this means that

$$\frac{1}{2}F^{\mu\nu}F_{\mu\nu} = \left(\mathbf{B}^2 - \frac{\mathbf{E}^2}{c^2}\right) = \left(\mathbf{B}^2 + \frac{\mathbf{E}_{\mathrm{E}}^2}{c^2}\right). \tag{3.92}$$

Notice how the Euclideanized Lagrangian takes on the appearance of a Hamiltonian. This result is the key to relating Wick-rotated field theory to thermodynamical partition functions. It works because the quantum phase factor $\exp(\mathrm{i}S_{\mathrm{M}}/\hbar)$ looks like the partition function, or statistical weight factor $\exp(-\beta H_{\mathrm{M}})$ when Wick-rotated:

$$S_{\mathrm{E}} = -\mathrm{i}S_{\mathrm{M}}, \tag{3.93}$$

since the volume measure $\mathrm{d}V_x^{\mathrm{E}} = -\mathrm{i}\mathrm{d}V_x$. The superficial form of the Lagrangian density is unchanged in theories with only quadratic derivatives provided everything is written in terms of summed indices, but internally all of the time-summed terms have changed sign. Thus, one has that

$$\exp\left(\mathrm{i}\frac{S_{\mathrm{M}}}{\hbar}\right) = \exp\left(-\frac{S_{\mathrm{E}}}{\hbar}\right) \sim \exp\left(-\frac{1}{\hbar}\int \mathrm{d}V_{\mathrm{E}}\ \mathcal{H}_{\mathrm{M}}\right). \tag{3.94}$$

A Euclideanized invariant becomes something which looks like a Minkowski space non-invariant. The invariant F^2, which is used to deduce the dynamics of electromagnetism, transformed into Euclidean space, resembles a non-invariant of Minkowski space called the *Hamiltonian*, or total energy function (see eqn. (2.70)). This has physical as well as practical implications for field theories at finite temperature. If one takes the Euclidean time to be an integral from zero to $\hbar\beta$ and take $H = \int \mathrm{d}\sigma\mathcal{H}$,

$$\exp\left(\mathrm{i}\frac{S_{\mathrm{M}}}{\hbar}\right) = \exp\left(-\frac{1}{\beta}H_{\mathrm{M}}\right), \tag{3.95}$$

then a Euclidean field theory phase factor resembles a Minkowski space, finite-temperature Boltzmann factor. This is discussed further in chapter 6.

In a Cartesian basis, one has

$$g_{\mu\nu} \to g_{\mu\nu}^{\mathrm{E}} = \delta_{\mu\nu}. \tag{3.96}$$

https://doi.org/10.1017/9781009289887.005 Published online by Cambridge University Press