

10

Kinematical and dynamical transformations

In addition to parameter symmetries, which express geometrical uniformity in spacetime, some symmetries relate to uniformities in the more abstract space of the dynamical variables themselves. These ‘internal’ symmetries can contain group elements which depend on the spacetime parameters, so that there is a cross-dependence on the internal and external parameters; they are intimately connected to the concept of ‘charge’ (see also chapter 12).

Internal symmetries are not necessarily divorced from geometrical (parameter) invariances, but they may be formulated independently of them. The link between the two is forged by the spacetime properties of the action principle, through interactions between fields which act as generators for the symmetries (see, for instance, section 11.5).

10.1 Global or rigid symmetries

The simplest symmetries are *global symmetries*, whose properties are independent of spacetime location. For example, the action

$$S = \int (dx) \left\{ \frac{1}{2} (\partial^\mu \phi)(\partial_\mu \phi) + \frac{1}{2} m^2 \phi^2 \right\} \quad (10.1)$$

is invariant under the Z_2 reflection symmetry $\phi(x) \rightarrow -\phi(x)$ at all spacetime points. This symmetry would be broken by a term of the form

$$S = \int (dx) \left\{ \frac{1}{2} (\partial^\mu \phi)(\partial_\mu \phi) + \frac{1}{2} m^2 \phi^2 + \frac{1}{3!} \alpha \phi^3 \right\}. \quad (10.2)$$

The next most commonly identified symmetry is the $U(1)$ phase symmetry, which is exhibited by complex fields:

$$\Phi \rightarrow e^{i\theta} \Phi. \quad (10.3)$$

The action

$$S = \int (dx) \left\{ \frac{1}{2} (\partial^\mu \Phi^*) (\partial_\mu \Phi) + \frac{1}{2} m^2 \Phi^* \Phi \right\} \quad (10.4)$$

is invariant under this transformation. This symmetry is related to the idea of electric charge. One can say that charge is a book-keeping parameter for this underlying symmetry, or vice versa.

Multi-component fields also possess global symmetries. For instance, the model

$$S = \int (dx) \left\{ \frac{1}{2} (\partial^\mu \phi_A) (\partial_\mu \phi_A) + \frac{1}{2} m^2 \phi_A \phi_A \right\} \quad (10.5)$$

is invariant under the transformation

$$\phi_A = U_A^B \phi_B, \quad (10.6)$$

where

$$U_A^B U_B^C = \delta_A^C, \quad (10.7)$$

or $U^T U = I$. This is the group of orthogonal transformations $O(N)$, where $A, B = 1, \dots, N$. Multi-level atom bound states can be represented in this way, see, for instance, section 10.6.3. Multi-component symmetries of this kind are form groups which are generally *non-Abelian* (see chapter 23 for further details on the formulation of non-Abelian field theory).

The physical significance of global symmetries is not always clear *a priori*. They represent global correlations of properties over the whole of spacetime simultaneously, which apparently contradicts special relativity. Often the analysis of global symmetries is only a prelude to studying local ones. Even in section 10.6.3, the global symmetry appears only as a special case of a larger local symmetry. One often finds connections between spacetime symmetries and phase symmetries which make local symmetries more natural. This is especially true in curved spacetime or inhomogeneous systems.

In practice, global symmetries are mainly used in non-relativistic, small systems where simultaneity is not an issue, but there is a lingering suspicion that global symmetries are only approximations to more complex local ones.

10.2 Local symmetries

A symmetry is called *local* if it involves transformations which depend on coordinates. Allowing a phase transformation to depend on the coordinates is sometimes referred to as ‘gauging the symmetry’. For example, the local version of the complex $U(1)$ symmetry is

$$\begin{aligned} \Phi &\rightarrow e^{i\theta(x)} \phi \\ \Gamma_\mu(x) &\rightarrow \Gamma_\mu - (\partial_\mu \theta). \end{aligned} \quad (10.8)$$

The action now needs to be modified in order to account for the fact that partial derivatives do not commute with these transformations. The partial derivative is exchanged for a covariant one, which includes the connection $\Gamma_\mu(x)$,

$$D_\mu = \partial_\mu + i\Gamma_\mu. \quad (10.9)$$

$$S = \int (dx) \left\{ \frac{1}{2} (D^\mu \Phi^*) (D_\mu \Phi) + \frac{1}{2} m^2 \Phi^* \Phi \right\}. \quad (10.10)$$

The most important way in which abstract field symmetries connect with spacetime properties is through the derivative operator, since this is the generator of dynamical behaviour in continuous, holonomic systems.

10.3 Derivatives with a physical interpretation

Covariance with respect to local symmetries of the action may be made manifest by re-writing the action in terms of an effective derivative. The physical motivation for this procedure is that the ordinary partial derivative does not have an invariant physical interpretation under local symmetry transformations. By adding additional terms, called ‘connections’, to a partial derivative ∂_μ , one creates an ‘effective derivative’, D_μ , which does have an invariant meaning. Although the definition of a new object, D_μ , is essentially a notational matter, the notation is important because it assigns a unique interpretation to the new derivative symbol, in any basis. For that reason, D_μ is called a covariant derivative.

There are two related issues in defining derivatives which have a physical interpretation. The first issue has to do with the physical assumption that measurable quantities are associated with Hermitian operators (Hermitian operators have real eigenvalues). The second has to do with form invariance under specific transformations.

10.3.1 Hermiticity

According to the standard interpretation of quantum mechanics, physical quantities are derived from Hermitian operators, since Hermitian operators have real eigenvalues. Hermitian operators are self-adjoint with respect to the scalar product:

$$(\phi | \mathcal{O} | \phi) = (\mathcal{O}^\dagger \phi, \phi) = (\phi, \mathcal{O} \phi), \quad (10.11)$$

or formally

$$\mathcal{O}^\dagger = \mathcal{O}. \quad (10.12)$$

If the operator \mathcal{O} is a derivative operator, it can be moved from the left hand side of the inner product to the right hand side and back by partial integration. This follows from the definition of the inner product. For example, in the case of the Schrödinger field, we have

$$\begin{aligned} (\psi_1, i\partial_\mu \psi_2) &= \int d\sigma \psi_1^\dagger (-i\partial_\mu \psi_2) \\ &= \int d\sigma (-i\partial_\mu \psi_1^\dagger) \psi_2 \\ &= -(i\partial_\mu \psi_1, \psi_2). \end{aligned} \tag{10.13}$$

Partial integration moves the derivative from ψ_2 to ψ_1 and changes the sign. This sign change means that $i\partial_\mu$ is not a Hermitian operator. In order for a derivative operator to be Hermitian, it must not change sign. Thus, a quadratic derivative, ∂^2 , would be Hermitian. For linear derivatives, we should symmetrize the left–right nature of the derivative. Using arrow notation to show the direction in which the derivative acts, we may write

$$i\partial_\mu \rightarrow \frac{i}{2}(\overset{\rightarrow}{\partial}_\mu - \overset{\leftarrow}{\partial}_\mu) \equiv \frac{i}{2} \overset{\leftrightarrow}{\partial}_\mu. \tag{10.14}$$

Partial integration preserves the sign of $\overset{\leftrightarrow}{\partial}_\mu$.

A second important situation occurs when this straightforward partial integration is obstructed by a multiplying function. This is commonly the situation for actions in curvilinear coordinates where the Jacobian in the volume measure is a function of the coordinates themselves. The same thing occurs in momentum space. To see this, we note that the volume measure in the inner product is

$$d\sigma = |J(x)|d^n x, \tag{10.15}$$

where $J(x)$ is the Jacobian of the coordinates relative to a Cartesian basis. Normally, $J(x) = \sqrt{g_{ij}(x)}$, where $g_{ij}(x)$ is the spatial metric. If we now try to integrate by parts with this volume measure, we pick up an extra term involving the derivative of this function:

$$\int d\sigma \psi_1^\dagger (-i\partial_\mu \psi_2) = \int d\sigma \left(-i\partial_\mu - i \frac{\partial_\mu J}{J} \right) \psi_1^\dagger \psi_2. \tag{10.16}$$

This problem affects x derivatives in curved x coordinates and k derivatives in Fourier transform space, on the ‘mass shell’. See table 10.1.

The partial derivatives in table 10.1 are clearly not Hermitian. The problem now is the extra term owing to the coordinate-dependent measure. We can solve this problem by introducing an extra term, called a ‘connection’, which makes the derivative have the right properties under integration by parts. The crux of the matter is to find a linear derivative operator which changes sign under

Table 10.1. Derivatives and measures.

Derivative	Measure
∂_μ	$\sqrt{g_{ij}(x)}d^n \mathbf{x}$
$\frac{\partial}{\partial k_\mu}$	$\frac{d^n \mathbf{k}}{2\omega(k)}$

integration by parts, but does not pick up any new terms. Then we are back to the first example above, and further symmetrization is trivial. Consider the spacetime derivative. The problem will be solved if we define a new derivative by

$$D_\mu = \partial_\mu + \Gamma_\mu, \quad (10.17)$$

and demand that Γ_μ be determined by requiring that D_μ only change sign under partial integration:

$$\int d^n \mathbf{x} J(x) \phi_1(D_\mu \phi_2) = \int d^n \mathbf{x} J(x) (-D_\mu \phi_1) \phi_2. \quad (10.18)$$

Substituting eqn. (10.17) into eqn. (10.18), we find that Γ_μ must satisfy

$$-(\partial_\mu J) + M\Gamma_\mu = -M\Gamma_\mu, \quad (10.19)$$

or

$$\Gamma_\mu = \frac{1}{2} \frac{\partial_\mu J}{J}. \quad (10.20)$$

The new derivative D_μ can be used to construct symmetrical derivatives such as $D^2 = D_\mu D^\mu$ and $\overleftrightarrow{D}_\mu$, by analogy with the partial derivative.

10.3.2 Commutativity with transformations

The problem of additional terms arising due to the presence of functions of the coordinates occurs not just with the integration measure but also with transformations of the fields. Imagine a field theory involving the field variable $\phi(x)$, a simple scalar field satisfying an equation of motion given by

$$-\square \phi = -\partial_\mu \partial^\mu \phi = 0. \quad (10.21)$$

We then consider the transformation

$$\phi(x) \rightarrow \phi(x)U(x), \quad (10.22)$$

where $U(x)$ is an arbitrary function of x . This situation crops up quite often in field theory, when $U(x)$ is a phase transformation. The first thing we notice is that our equation of motion (10.21) is neither covariant nor invariant under this transformation, since

$$\partial_\mu \phi \rightarrow (\partial_\mu \phi(x))U(x) + (\partial_\mu U(x))\phi(x). \quad (10.23)$$

Clearly eqn. (10.21) is only a special case of the equations of motion. Under a transformation we will always pick up new terms, as in eqn. (10.23), since the partial derivative does not commute with an arbitrary function $U(x)$, so $U(x)$ can never be cancelled out of the equations. But, suppose we re-write eqn. (10.23) as

$$(\partial_\mu \phi(x))U(x) = U(x) \left(\partial_\mu + \frac{\partial_\mu U}{U} \right) \phi(x), \quad (10.24)$$

and define a new derivative

$$D_\mu = (\partial_\mu + \Gamma_\mu), \quad (10.25)$$

where $\Gamma_\mu = U^{-1}(\partial_\mu U) = \partial_\mu \ln U$, then we have

$$\partial_\mu (U(x)\phi(x)) = U(x)D_\mu(\phi(x)). \quad (10.26)$$

We can now try to make eqn. (10.21) covariant. We replace the partial derivative by a covariant one, giving

$$-\partial^2 \phi(x) = -D_\mu D^\mu \phi(x) = 0. \quad (10.27)$$

The covariance can be checked by applying the transformation

$$-D^2(U(x)\phi(x)) = -U(x)\partial^2(\phi(x)) = 0 \quad (10.28)$$

so that the factor of $U(x)$ can now be cancelled from both sides.

At this point, it almost looks as though we have achieved an invariance in the form of the equations, but that is not so. To begin with, the derivative we introduced only works for a specific function $U(x)$, and that function is actually buried in the definition of the new derivative, so all we have done is to re-write the equation in a new notation. If we change the function, we must also change the derivative. Also, if we add a source to the right hand side of the equations, then this argument breaks down. In other words, while the equation is now written in a more elegant way, it is neither covariant nor invariant since the specific values of the terms must still change from case to case.

10.3.3 Form-invariant derivatives

To obtain invariance requires another idea – and this involves a physical assumption. Instead of defining $\Gamma_\mu = U^{-1}(\partial_\mu U)$, we say that Γ_μ is itself a new

physical field; in addition, we demand that the transformation law be *extended* to include a transformation of the new field Γ_μ . The new transformation rule is then

$$\begin{aligned}\phi(x) &\rightarrow U(x)\phi(x) \\ \Gamma_\mu &\rightarrow \Gamma_\mu - \frac{\partial_\mu f}{f}.\end{aligned}\quad (10.29)$$

Γ_μ might be zero in some basis, but not always. Under this new assumption, only the physical fields transform. The covariant derivative is form-invariant, as are the equations of motion, since Γ_μ absorbs the extra term which is picked up by the partial differentiation.

Note how this last step is a physical assumption. Whereas everything leading up to eqn. (10.28) has simply been a mathematical manipulation of the formulation, the assumption that Γ_μ is a new field, which transforms separately, is a physical assumption. This makes symmetries of this type *dynamical symmetries*, rather than coincidental kinematical symmetries, which arise simply as a matter of fortuitous cancellations.

The covariant derivative crops up in several guises – most immediately in connection with the interaction of matter with the electromagnetic field, and the invariance of probabilities under arbitrary choice of quantum phase.

10.4 Charge conjugation

A charge conjugation transformation, for a field with sufficient internal symmetry, is defined to be one which has the following properties on spin 0, $\frac{1}{2}$, and 1 fields:

$$\begin{aligned}\mathcal{C} \phi(x) \mathcal{C}^\dagger &= \eta_\phi \phi^\dagger(x) \\ \mathcal{C} \psi(x) \mathcal{C}^\dagger &= \eta_\psi \bar{\psi}^T(x) \\ \mathcal{C} A_\mu(x) \mathcal{C}^\dagger &= -A_\mu.\end{aligned}\quad (10.30)$$

Under this transformation, the sign of the gauge field (and hence the sign of the charge it represents) is reversed. It is clearly a discrete rather than a continuous transformation. In the complex scalar case, the transformation simply exchanges the conjugate pair of fields. This is easy to see in the formulation of the complex scalar as a pair of real fields (see section 19.7), where the field, A_μ , is accompanied by the anti-symmetric tensor ϵ_{AB} , which clearly changes sign on interchange of scalar field components. In the Dirac spinor case, a more complicated transformation is dictated by the Dirac matrices (see section 20.3.4).

10.5 TCP invariance

The TCP theorem [87, 88, 105, 114] asserts that any local physical Lagrangian must be invariant under the combined action of time reversal (T), parity (P) and charge conjugation (C). More specifically, it claims that the effect of CP should be the same as T. Interactions may be constructed which violate these symmetries separately, but the TCP theorem requires the product of these transformations

$$U_{TCP} = U_T U_P U_C \tag{10.31}$$

to be conserved:

$$\begin{aligned} U_{TCP} \phi(x) U_{TCP}^{-1} &= \eta_c \eta_t \eta_p \phi^\dagger(-x) \\ U_{TCP} \psi(x) U_{TCP}^{-1} &= -\gamma_5 \eta_c \eta_t \eta_p \psi^*(-x) \\ U_{TCP} A_\mu(x) U_{TCP}^{-1} &= -\eta_c \eta_t \eta_p A_\mu^\dagger(-x). \end{aligned} \tag{10.32}$$

A choice of phase such that $\eta_c \eta_t \eta_p = 1$ is natural. This transformation has particularly interesting consequence in the case of a spin- $\frac{1}{2}$ field. If one considers a bi-linear term in the action, of the form

$$\Delta\mathcal{L} = \bar{\psi}_1(x) \mathcal{O} \psi_2(x), \tag{10.33}$$

then the application of the transformation leads to

$$\begin{aligned} U_{TCP} [\bar{\psi}_1(x) \mathcal{O}(x) \psi_2(x)] U_{TCP}^{-1} &= U_{TCP} [\psi_1^\dagger \gamma^0 \mathcal{O}(x) \psi_2(x)] U_{TCP}^{-1} \\ &= [\psi_1(-x)^\dagger \gamma_5 \gamma^0 \mathcal{O}(x) \gamma_5 \psi_2(-x)] \\ &= -[\bar{\psi}_1^\dagger(-x) \gamma_5 \mathcal{O}(x) \gamma_5 \psi_2(-x)] \\ &= [\bar{\psi}_1^\dagger(-x) \gamma_5 \mathcal{O}(x) \gamma_5 \psi_2(-x)]^\dagger. \end{aligned} \tag{10.34}$$

In the last two lines, a minus sign appears first when commuting γ_5 through γ^0 , then a second minus sign must be associated with commuting $\bar{\psi}_1$ and ψ_2 . Under the combination of TCP, one also has scalar behaviour

$$\gamma_5 \mathcal{O}(x) \gamma_5 = -\mathcal{O}(-x). \tag{10.35}$$

Regardless of what one chooses to view as fundamental, the invariance under TCP and the anti-commutativity of the Dirac field go hand in hand

$$U_{TCP} [\bar{\psi}_1(x) \mathcal{O}(x) \psi_2(x)] U_{TCP}^{-1} = [\bar{\psi}_1^\dagger(-x) \mathcal{O}(-x) \psi_2(-x)]^\dagger. \tag{10.36}$$

What is noteworthy about the TCP theorem is that it relates environmental, spacetime symmetries (space and time reflection) to internal degrees of freedom (charge reflection). This result follows from the locality and Hermiticity of the action, but requires also a new result: the spin-statistics theorem, namely that spin- $\frac{1}{2}$ particles must anti-commute. This means that fermionic variables should be represented by anti-commuting Grassman variables.

10.6 Examples

The following examples show how symmetry requirements and covariance determine the structure of the action under both internal and spacetime symmetries. The link between spacetime and internal symmetry, brought markedly to bear in the TCP theorem, is also reflected through conformal symmetry and transformation connections.

10.6.1 Gauge invariance: electromagnetism

The Schrödinger equation has the form

$$\left(-\frac{\hbar^2}{2m}\partial^i\partial_i + V\right)\psi = i\partial_t\psi. \quad (10.37)$$

The wavefunction $\psi(x)$ is not a direct physical observable of this equation. However, the probability

$$P = |\psi|^2 \quad (10.38)$$

is observable. As the modulus of a complex number, the probability is invariant under phase transformations of the form

$$\psi(x) \rightarrow e^{i\theta(x)}\psi(x). \quad (10.39)$$

One expects that the Schrödinger action should be invariant under this symmetry too. It should be clear from the discussion in section 10.3 that this is not the case as long as the phase $\theta(x)$ is x -dependent; to make the Schrödinger equation invariant, we must introduce a new field, A_μ . By appealing to the phenomenology of the Aharonov–Bohm effect, one can identify A_μ with the electromagnetic vector potential.

From eqn. (2.44), one may assume the following form for the covariant derivative:

$$-i\hbar\partial_\mu \rightarrow -i\hbar D_\mu = -i\hbar\left(\partial_\mu - i\frac{e}{\hbar}A_\mu\right), \quad (10.40)$$

since it only differs from a completely general expression by some constants c , \hbar and e . In explicit terms, we have chosen $\Gamma_\mu = -i\frac{e}{\hbar}A_\mu$. The total gauge or phase transformation is now a combination of eqns. (10.37) and (10.39), and to secure invariance of the equation, we must perform both transformations together.

Applying the phase transformation and demanding that D_μ commute with the phase leads to

$$\begin{aligned} D_\mu(e^{i\theta(x)}\psi(x)) &= e^{i\theta(x)}\left(\left(\partial_\mu - i\frac{e}{\hbar}(A_\mu + \partial_\mu\theta)\right) + i(\partial_\mu\theta)\right)\psi(x), \\ &= e^{i\theta(x)}D_\mu(\psi(x)) \end{aligned} \quad (10.41)$$

where $D_\mu = \partial_\mu - i\frac{e}{\hbar}A_\mu$, and the last line follows provided we take

$$i(\partial_\mu\theta) - i\frac{e}{\hbar}(\partial_\mu s) = 0. \quad (10.42)$$

Both $\theta(x)$ and $s(x)$ are completely arbitrary scalar fields, so this relation merely identifies them to be the same arbitrary quantity. We may therefore write the combined phase and gauge transformations in the final form

$$\begin{aligned} \psi(x) &\rightarrow \psi'(x) = e^{i\frac{e}{\hbar}s(x)}\psi(x) \\ A_\mu(x) &\rightarrow A'_\mu(x) = A_\mu(x) + (\partial_\mu s(x)), \end{aligned} \quad (10.43)$$

and Schrödinger's equation in gauge-invariant form is

$$\left(-\frac{\hbar^2}{2m}D^i D_i + V\right)\psi(x) = i\hbar D_t\psi, \quad (10.44)$$

where $D_t = cD_0$. In terms of the covariant derivative, we can write the field strength tensor as a commutator:

$$[D_\mu, D_\nu] = -2i\frac{e}{\hbar}F_{\mu\nu}. \quad (10.45)$$

This may be compared with eqn. (10.58) in the following section.

10.6.2 Lorentz invariance: gravity

In the presence of a non-trivial metric $g_{\mu\nu}$, i.e. in the curved spacetime of a gravitational field, or in a curvilinear coordinate system, the Lorentz transformation is not merely a passive kinematic transformation, it has the appearance of a dynamical transformation. This change of character is accompanied by the need for a transforming connection, like the ones above, only now using a more complex rule, fit for general tensor fields.

The Lorentz-covariant derivative is usually written ∇_μ , so that covariance is obtained by substituting partial derivatives in the following manner:

$$\partial_\mu \rightarrow \nabla_\mu. \quad (10.46)$$

With Lorentz transformations there is a subtlety, since we are interested in many different representations of the Lorentz group, i.e. in tensors of different rank. For scalar fields, there is no problem for Lorentz transformations. A scalar field does not transform under a Lorentz transformation, so the partial derivative is Hermitian. In other words,

$$\nabla_\mu\phi(x) = \partial_\mu\phi(x). \quad (10.47)$$

For a vector field, however, the story is different. Now the problem is that vectors transform according to the rules of tensor transformations and the partial derivative of a vector field does not commute with Lorentz transformations themselves. To fix this, a connection is required.¹ As before, we look for a connection which makes the derivative commute with the transformation law. Consider the vector field V_μ . Let us transform it from one set of coordinates, ξ^α, ξ^β , to another, x^μ, x^ν . According to the rules of tensor transformation, we have

$$\begin{aligned} V'_\mu(\xi) &= \frac{\partial \xi^\beta}{\partial x^\mu} V_\beta(x) \\ &= (\partial_\mu^x \xi^\beta) V_\beta(x) \\ &= L_\mu^\beta V_\beta(x). \end{aligned} \tag{10.48}$$

Let us now introduce a derivative ∇_μ with the property that

$$\nabla(LV) = L(\nabla'V), \tag{10.49}$$

i.e. such that the derivative ∇_μ is form-invariant, but transforms dynamically under a coordinate transformation. Let us write

$$\nabla_\mu = \partial_\mu + \Gamma_{\mu?}, \tag{10.50}$$

where the question mark is to be determined. At this stage, it is not clear just how the indices will be arranged on Γ , since there are several possibilities when acting on a vector field. Let us evaluate

$$\begin{aligned} \nabla_\mu V'_\nu(x) &= \nabla_\mu (L_\nu^\beta V_\beta) \\ &= \nabla_\mu \left((\partial_\nu^x \xi^\beta(x)) V_\beta \right) \\ &= (\partial_\mu + \Gamma_\mu) \left((\partial_\nu^x \xi^\beta(x)) V_\beta \right) \\ &= (\partial_\mu \partial_\nu \xi^\beta) V_\beta(x) + (\partial_\nu \xi^\beta)(\partial_\mu V_\beta) + \Gamma_\mu (\partial_\nu \xi^\beta) V_\beta. \end{aligned} \tag{10.51}$$

From the assumed form of the transformation, we expect this to be

$$L_\nu^\beta (\nabla'_\mu V_\beta) = (\partial_\mu^x \xi^\beta)(\partial_\nu V_\beta) + \Gamma'_\mu V_\beta. \tag{10.52}$$

Comparing eqn. (10.51) and eqn. (10.52), we see that

$$(\partial_\mu^x \xi^\beta) \Gamma_\mu \rightarrow (\partial_\mu^x \xi^\beta) \Gamma'_\mu - (\partial_\mu \partial_\nu \xi^\beta). \tag{10.53}$$

¹ There are two ways to derive the connection for Lorentz transformations, one is to look at the Hermitian nature of the derivatives; the other is to demand that the derivative of a tensor always be a tensor. Either way, one arrives at the same answer, for essentially the same reason.

Multiplying through by $(\partial_\alpha^\xi x^\nu)$ and using the chain-rule, we see that the transformation of Γ must be

$$\Gamma \rightarrow \Gamma' - (\partial_\alpha^\xi x^\nu)(\partial_\mu^x \partial_\nu^x \xi^\beta). \tag{10.54}$$

This also shows us that there must be three indices on Γ , so that the correct formulation of the vector-covariant derivative is

$$\nabla_\mu V_\nu = \partial_\mu V_\nu - \Gamma_{\mu\nu}^\lambda V_\lambda, \tag{10.55}$$

with transformation rule

$$\Gamma_{\alpha\mu}^\beta \rightarrow \Gamma'_{\alpha\mu}{}^\beta - (\partial_\alpha^\xi x^\nu)(\partial_\mu^x \partial_\nu^x \xi^\beta). \tag{10.56}$$

Thus, demanding commutativity with a dynamical transformation, once again requires the introduction of a corrective term, or connection.

What turns a coordinate transformation into a dynamical transformation is the spacetime dependence of the metric. It makes the coordinate transformation into a spacetime-dependent quantity also, changing its status from a passive kinematical property to an active dynamical one. The non-linearity which is implied by having coordinates which depend on other coordinates is what leads Einstein’s theory of gravity to use the concept of intrinsic curvature.

The above procedure can be generalized to any tensor field. Extra terms will be picked up for each index, since there is a coordinate transformation term for each index of a tensor. The sign of the correction depends on whether indices are raised or lowered, because of the mutually reciprocal nature of the transformations in these cases. To summarize, we have spacetime-covariant derivatives defined as follows:

$$\begin{aligned} \nabla_\mu \phi(x) &= \partial_\mu \phi(x) \\ \nabla_\mu A_\nu &= \partial_\mu A_\nu - \Gamma_{\mu\nu}^\lambda A_\lambda \\ \nabla_\mu A^\nu &= \partial_\mu A^\nu + \Gamma_{\mu\lambda}^\nu A^\lambda \\ \nabla_\mu T_{\lambda}^{\mu\sigma} &= \partial_\mu T_{\lambda}^{\mu\sigma} + \Gamma_{\rho\nu}^\mu T_{\lambda}^{\nu\sigma} + \Gamma_{\rho\nu}^\sigma T_{\lambda}^{\mu\nu} - \Gamma^\kappa T_{\kappa}^{\mu\sigma}. \end{aligned} \tag{10.57}$$

Note that we can express the curvature as a commutator of covariant derivatives:

$$[\nabla_\mu, \nabla_\nu] \xi^\sigma = -R^\lambda_{\sigma\mu\nu} \xi^\lambda. \tag{10.58}$$

This may be compared with eqn. (10.45).

10.6.3 The two-level atom in a strong radiation field

It was first realized by Jaynes and Cummings that a semi-classical model of a two-level atom could reproduce the essential features of the quantum theoretical

problem [79]. The two-level system has a broad repertoire of applications in physics, from spin models to the micromaser [91]. It is related to a class of Dicke models [37, 57], and, in the so-called rotating wave approximation, it becomes the Jaynes–Cummings model [79] which may be solved exactly. A Hamiltonian analysis of symmetries in this Jaynes–Cummings model is given in ref. [7].

The symmetry techniques and principles of covariant field theory can be applied to the two-level atom to solve the full model and eliminate the need for the so-called rotating wave approximation. Consider the phenomenological two-level system described by the action

$$S = \int (dx) \left[-\frac{\hbar^2}{2m} (\partial^i \psi_A)^* (\partial_i \psi_A) - \psi_A^* V_{AB}(t) \psi_B + \frac{i\hbar}{2} (\psi^* D_t \psi - (D_t \psi)^* \psi) \right], \quad (10.59)$$

where $A, B = 1, 2$ characterizes the two levels, $i\hbar D_t = i\hbar \partial_t + i\Gamma(t)$ in matrix notation, and $\Gamma = \Gamma_{AB}$ is an off-diagonal anti-symmetrical matrix. At frequencies which are small compared with the light-size of the atom, an atom may be considered electrically neutral. The distribution of charge within the atoms is not required here. In this approximation the leading interaction is a resonant dipole transition. The connection Γ_{AB} plays an analogous role to the electromagnetic vector potential in electrodynamics, but it possesses no dynamics of its own. Rather, it works as a constraint variable, or auxiliary Lagrange multiplier field. There is no electromagnetic vector potential in the action, since the field is electrically neutral in this formulation. Γ_{AB} refers not to the $U(1)$ phase symmetry but to the two-level symmetry. Variation of the action with respect to $\Gamma(t)$ provides us with the conserved current.

$$\frac{\delta S}{\delta \Gamma_{AB}} = \frac{i}{2} (\psi_A^* \psi_B - \psi_B^* \psi_A), \quad (10.60)$$

which represents the amplitude for stimulated transition between the levels. The current generated by this connection is conserved only on average, since we are not taking into account any back-reaction. The conservation law corresponds merely to

$$\partial_t \left(\frac{\delta S}{\delta \Gamma_{AB}} \right) \propto \sin \left(2 \int X(t) \right), \quad (10.61)$$

where $X(t)$ will be defined later. The potential $V_{AB}(t)$ is time-dependent, and comprises the effect of the level splitting as well as a perturbation mediated by the radiation field. A ‘connection’ $\Gamma_{21} = -\Gamma_{12}$ is introduced since the diagonalization procedure requires a time-dependent unitary transformation,

and thus general covariance demands that this will transform in a different basis. The physics of the model depends on the initial value of this ‘connection’, and this is the key to the trivial solubility of the Jaynes–Cummings model.

In matrix form we may write the action for the matter fields

$$S = \int (dx) \psi_A^* \mathcal{O}_{AB} \psi_B \tag{10.62}$$

where

$$\mathcal{O} = \begin{bmatrix} -\frac{\hbar^2 \nabla^2}{2m} - V_1 - \frac{i\hbar}{2} \overleftrightarrow{D}_t & J(t) + i\Gamma_{12} \\ J(t) - i\Gamma_{12} & -\frac{\hbar^2 \nabla^2}{2m} - V_2 - \frac{i\hbar}{2} \overleftrightarrow{D}_t \end{bmatrix}. \tag{10.63}$$

The level potentials may be regarded as constants in the effective theory. They are given by $V_1 = E_1$ and $V_2 = E_2 - \hbar\Omega_R$ where $\hbar\Omega_R$ is the interaction energy imparted by the photon during the transition, i.e. the continuous radiation pressure on the atom. In the effective theory, we must add this by hand, since we have separated the levels into independent fields which are electrically neutral; it would follow automatically in a complete microscopic theory. The quantum content of this model is now that this recoil energy is a quantized unit of $\hbar\Omega$, the energy of a photon at the frequency of the source. Also, the amplitude of the source, J , would be quantized and proportional to the number of photons on the field. If one switches off the source (which models the photon’s electric field), this radiation energy does not automatically go to zero, so this form is applicable mainly to continuous operation (stimulation). The origin of the recoil is clear, however: it is the electromagnetic force’s interaction with the electron, transmitted to the nucleus by binding forces. What we are approximating is clearly a $J^\mu A_\mu$ term for the electron, with neutralizing background charge.

It is now desirable to perform a unitary transformation on the action $\psi \rightarrow U\psi$, $\mathcal{O} \rightarrow U\mathcal{O}U^{-1}$, which diagonalizes the operator \mathcal{O} . Clearly, the connection Γ_{AB} will transform under this procedure by

$$\Gamma \rightarrow \Gamma + \frac{i\hbar}{2} (U(\partial_t U^{-1}) - (\partial_t U)U^{-1}) \tag{10.64}$$

since a time-dependent transformation is required to effect the diagonalization.

For notational simplicity we define $\hat{L} = -\frac{\hbar^2 \nabla^2}{2m} - \frac{i\hbar}{2} \overleftrightarrow{D}_t$, so that the secular equation for the action is:

$$(\hat{L} - E_1 - \lambda)(\hat{L} - E_2 + \hbar\Omega - \lambda) - (J^2 + \Gamma_{12}^2) = 0. \tag{10.65}$$

Note that since $J \overleftrightarrow{\partial}_t J = 0$ there are no operator difficulties with this equation. The eigenvalues are thus

$$\lambda_{\pm} = \hat{L} - \bar{E}_{12} + \hbar\Omega \pm \sqrt{\frac{1}{4}(\bar{E}_{21} - \hbar\Omega)^2 + J^2 + \Gamma_{12}^2} \tag{10.66}$$

$$\equiv \hat{L} - \bar{E}_{12} + \hbar\Omega \pm \sqrt{\hbar^2\tilde{\omega}^2 + J^2 + \Gamma_{12}^2} \tag{10.67}$$

$$\equiv \hat{L} - \bar{E}_{12} + \hbar\Omega \pm \hbar\omega_R, \tag{10.68}$$

where $\bar{E}_{12} = \frac{1}{2}(E_1 + E_2)$ and $\tilde{E}_{21} = (E_2 - E_1)$. For notational simplicity we define $\tilde{\omega}$ and ω_R . One may now confirm this procedure by looking for the eigenvectors and constructing U^{-1} as the matrix of these eigenvectors. This may be written in the form

$$U^{-1} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \tag{10.69}$$

where

$$\cos \theta = \frac{\hbar(\tilde{\omega} + \omega_R)}{\sqrt{\hbar^2(\tilde{\omega} + \omega_R)^2 + J^2 + \Gamma_{12}^2}} \tag{10.70}$$

$$\sin \theta = \frac{\sqrt{J^2 + \Gamma_{12}^2}}{\sqrt{\hbar^2(\tilde{\omega} + \omega_R)^2 + J^2 + \Gamma_{12}^2}}. \tag{10.71}$$

The change in the connection $\Gamma(t)$ is thus off-diagonal and anti-symmetric, as required by the gauge symmetry conservation law:

$$U \partial_t U^{-1} = \begin{pmatrix} 0 & \partial_t \theta \\ -\partial_t \theta & 0 \end{pmatrix}. \tag{10.72}$$

The time derivative of $\theta(t)$ may be written in one of two forms, which must agree

$$(\partial_t \theta) = \frac{\partial_t \cos \theta}{-\sin \theta} = \frac{\partial_t \sin \theta}{\cos \theta}. \tag{10.73}$$

This provides a consistency condition, which may be verified, and leads to the proof of the identities

$$\omega_R \partial_t \omega_R = J \partial_t J + \Gamma \partial_t \Gamma \tag{10.74}$$

and

$$\sqrt{J^2 + \Gamma^2}(\partial_t + \Lambda)\sqrt{J^2 + \Gamma^2} + (\tilde{\omega} + \omega_R)(\partial_t + \Lambda)(\tilde{\omega} + \omega_R) = 0 \tag{10.75}$$

for arbitrary $J(t)$ and $\Gamma(t)$, where

$$\Lambda = -\frac{1}{2} \frac{\partial_t ((\tilde{\omega} + \omega_R)^2 + J^2 + \Gamma^2)}{(\tilde{\omega} + \omega_R)^2 + J^2 + \Gamma^2}. \tag{10.76}$$

These relations are suggestive of a conformal nature to the transformation and, with a little manipulation using the identities, one evaluates

$$\Gamma_{12}/\hbar = (\partial_t \theta) = \frac{(J \partial_t J + \Gamma \partial_t \Gamma)}{\omega_R \sqrt{J^2 + \Gamma^2}} \left[1 - \frac{(\tilde{\omega} + \omega_R)(\tilde{\omega} + 2\omega_R)}{(\tilde{\omega} + \omega_R)^2 + J^2 + \Gamma^2} \right]. \tag{10.77}$$

This quantity vanishes when $J^2 + \Gamma^2$ is constant with respect to time. Owing to the various identities, the result presented here can be expressed in many equivalent forms. In particular, it is zero when $\tilde{\omega} = 0$. The equations of motion for the transformed fields are now

$$\begin{bmatrix} \hat{L} - \bar{E}_{12} + \hbar\omega_R & i\partial_t \theta \\ -i\partial_t \theta & \hat{L} - \bar{E}_{12} - \hbar\omega_R \end{bmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = 0. \tag{10.78}$$

In this basis, the centre of mass motion of the neutral atoms factorizes from the wavefunction, since a neutral atom in an electromagnetic field is free on average. The two equations in the matrix above may therefore be unravelled by introducing a ‘gauge transformation’, or ‘integrating factor’,

$$\psi_{\pm}(x) = e^{\pm i \int_0^t X(t') dt'} \bar{\psi}(x), \tag{10.79}$$

where the free wavefunction in $n = 3$ dimensions is

$$\bar{\psi}(x) = \int \frac{d\omega}{(2\pi)} \frac{d^n \mathbf{k}}{(2\pi)^n} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega t)} \delta(\chi) \tag{10.80}$$

is a general linear combination of plane waves satisfying the dispersion relation for centre of mass motion

$$\chi = \frac{\hbar^2 \mathbf{k}^2}{2m} + \hbar(\Omega - \omega) - \bar{E}_{12} = 0. \tag{10.81}$$

The latter is enforced by the delta function. This curious mixture of continuous (ω) and discontinuous (Ω) belies the effective nature of the model and the fact that its validity is only for a continuous operation (an eternally sinusoidal radiation source which never starts or stops). The relevance of the model is thus limited by this. Substituting this form, we identify $X(t)$ as the integrating factor for the uncoupled differential equations. The complete solution is therefore

$$\psi_{\pm}(x) = e^{\mp i \int_0^t (\omega_R + i\partial_t \theta) dt'} \bar{\psi}(x). \tag{10.82}$$

Notice that this result is an exact solution, in the sense of being in a closed form. In the language of a gauge theory this result is gauge-dependent. This is because our original theory was not invariant under time-dependent transformations. The covariant procedure we have applied is simply a method to transform the

equations into an appealing form; it does not imply invariance of the results under a wide class of sources.

That this system undergoes transitions in time may be seen by constructing wavefunctions which satisfy the boundary conditions where the probability of being in one definite state of the system is zero at $t = 0$. To this end we write $\Psi_1 = \frac{1}{2}(\psi_+ + \psi_-)$ and $\Psi_0 = \frac{1}{2i}(\psi_+ - \psi_-)$. In order to proceed beyond this point, it becomes necessary to specify the initial value of Γ_{12} . This choice carries with it physical consequences; the model is not invariant under this choice. The obvious first choice is to set this to zero. This would correspond to not making the rotating wave approximation in the usual two-level atom, with a cosine perturbation. Focusing on the state Ψ_0 which was unoccupied at $t = 0$ for $\Gamma_{12} = 0$,

$$\Psi_0 = \sin \left(\int_0^t dt' \left[\sqrt{\tilde{\omega}^2 + \hbar^{-2} J_0^2 \cos^2(\Omega t')} - i\tilde{\omega} \frac{J_0 \Omega \sin(\Omega t')}{2\hbar \omega_R} \left[\tilde{\omega} + \frac{J_0^2 \cos^2(\Omega t')}{\hbar^2 (\tilde{\omega} + \omega_R)} \right]^{-1} \right] \right) \bar{\psi}(x). \quad (10.83)$$

We are interested in the period, and the amplitude of this quantity, whose squared norm may be interpreted as the probability of finding the system in the prepared state, given that it was not there at $t = 0$. Although the integral is then difficult to perform exactly, it is possible to express it in terms of Jacobian elliptic integrals, logarithms and trig functions. Nevertheless it is clear that $\tilde{\omega} = \frac{1}{2}(\tilde{E}_{21}/\hbar - \Omega)$ is the decisive variable. When $\hbar\tilde{\omega} \ll J_0$ is small, the first term is $J_0 \cos(\Omega t)$ and the second term is small. This is resonance, although the form of the solution is perhaps unexpected. The form of the wavefunction guarantees a normalized result which is regular at $\tilde{\omega} = 0$, and one has $\Psi_0 \sim \sin \left(\int_0^t dt' \frac{J_0}{\hbar} \cos(\Omega t') \right)$, which may be compared with the standard result of the Jaynes–Cummings model $\Psi_0 \sim \sin(J_0 t/\hbar)$. In the quantum case the amplitude of the radiation source, J_0 , is quantized as an integral number, N_Ω , of photons of frequency Ω . Here we see modulation of the rate of oscillation by the photon frequency (or equivalently the level spacing). In a typical system, the photon frequency is several tens of orders of magnitude larger than the coupling strength $J_0 \ll \hbar\Omega \sim \tilde{E}_{12}$ and thus there is an extremely rapid modulation of the wavefunction. This results in an almost chaotic collapse–revival behaviour with no discernible pattern, far from the calm sinusoidal Rabi oscillations of the Jaynes–Cummings model. If $\hbar\tilde{\omega} \sim J_0$, the second term is of order unity, and then, defining the normalized resonant amplitude

$$A = \frac{J_0}{\sqrt{\hbar^2 \tilde{\omega}^2 + J_0^2}}, \quad (10.84)$$

one has

$$\Psi_0 \sim \sin \left(\frac{J_0 \Omega}{A} E(\Omega t, A) - A \int d(\Omega t) \frac{\sin(\Omega t)}{\sqrt{1 - A^2 \sin^2(\Omega t)}} \right) \bar{\psi}(x). \quad (10.85)$$

The Jacobian elliptical integral $E(\alpha, \beta)$ is a doubly periodic function, so one could expect qualitatively different behaviour away from resonance. On the other hand, far from resonance, $\hbar \tilde{\omega} \gg J_0$, the leading term of the connection becomes $\Psi_0 \sim \sin(\tilde{\omega} t) \bar{\psi}(x) \sim \sin(\Omega t) \bar{\psi}(x)$, and the effect of the level spacing is washed out.

One can also consider other values for the connection. Comparing Γ_{12} to the off-diagonal sources $\gamma^\mu D_\mu$, predicted on the basis of unitarity in effective non-equilibrium field theory [13], one obtains an indication that, if the initial connection is in phase with the time derivative of the perturbation, then one can effectively ‘re-sum’ the decay processes using the connection. This is a back-reaction effect of the time-dependent perturbation, or a renormalization in the language of ref. [13]. If one chooses $\Gamma_{12} = J_0 \sin(\Omega t)$, this has the effect of making the off-diagonal terms in the action not merely cosines but a complex conjugate pair $J_0 \exp(\pm i \Omega t)$. This corresponds to the result one obtains from making the rotating wave approximation near resonance. This initial configuration is extremely special. With this choice, one has exactly

$$\Psi_0 = \sin \left(\int_0^t dt' \left[\sqrt{\tilde{\omega}^2 + \hbar^{-2} J_0^2} \right] \right) \bar{\psi}(x). \quad (10.86)$$

The stability of the solution is noteworthy, and the diagonalizing transformation is rendered trivial. The connection $\partial_t \theta$ is now zero under the diagonalizing transformation. Thus, the above result is exact, and it is the standard result of the approximated Jaynes–Cummings model. This indicates that the validity of the Jaynes–Cummings model does not depend directly on its approximation, but rather on the implicit choice of a connection.

10.7 Global symmetry breaking²

The dynamical properties of certain interacting fields lead to solution surfaces whose stable minima favour field configurations, which are ordered, over random ones. Such fields are said to display the phenomenon of spontaneous ordering, or spontaneous symmetry breaking. This is a phenomenon in which the average behaviour of the field, in spite of all its fluctuations, is locked into a sub-set of its potential behaviour, with less symmetry. A classic example of

² $\hbar = c = \mu_0 = \epsilon_0 = 1$ in this section.

this is the alignment of spin in ferromagnetism, in which rotational symmetry is broken into a linear alignment.

Spontaneous symmetry breaking can be discussed entirely within the framework of classical field theory, but it should be noted that its dependence on interactions raises the problem of negative energies and probabilities, which is only fully resolved in the quantum theory of fields.

When a *continuous* global symmetry is broken (i.e. when its average state does not express the full global symmetry), one sees the appearance of massless modes associated with each suppressed symmetric degree of freedom. These massless modes are called Nambu–Goldstone bosons [59, 60, 99, 100]. To see how they arise, consider the action

$$S = \int (dx) \left\{ \frac{1}{2}(\partial^\mu \phi_A)(\partial_\mu \phi_A) + \frac{1}{2}m^2 \phi_A \phi_A + \frac{\lambda}{4!}(\phi_A \phi_A)^2 \right\}. \quad (10.87)$$

The interaction potential $V(\phi) = \frac{1}{2}m^2 \phi^2 + \frac{\lambda}{4!} \phi^4$ has a minimum at

$$ie\hbar c^2 \frac{\partial V(\phi)}{\partial \phi_A} = m^2 \phi_A + \frac{\lambda}{6} \phi_A (\phi_B \phi_B) = 0. \quad (10.88)$$

This would therefore be the equilibrium value for the average field. Note that a non-zero value for $\langle \phi \rangle$, within a bounded potential $\lambda > 0$, is possible only if $m^2 < 0$. Suppose one now considers the effect of fluctuations, or virtual processes, in the field. Following the procedure of chapter 6, one may split the field into an average (constant) part $\langle \phi \rangle$ and a fluctuating (quickly varying) part φ ,

$$\phi_A = \langle \phi \rangle_A + \varphi_A. \quad (10.89)$$

Expressed in terms of these parts, the terms of the action become:

$$\begin{aligned} (\partial^\mu \phi_A)(\partial_\mu \phi_A) &\rightarrow (\partial^\mu \varphi)(\partial_\mu \varphi) \\ \frac{1}{2}m^2 \phi_A \phi_A &\rightarrow \frac{1}{2}(\langle \phi \rangle_A \langle \phi \rangle_A + 2\langle \phi \rangle_A \varphi_A + \varphi_A \varphi_A) \\ (\phi_A \phi_A)^2 &\rightarrow (\langle \phi \rangle_A \langle \phi \rangle_A) + 4(\langle \phi \rangle_A \langle \phi \rangle_A)(\langle \phi \rangle_B \varphi_B) \\ &\quad + 2(\varphi_A \varphi_A)(\langle \phi \rangle_B \langle \phi \rangle_B) + 4(\varphi_A \langle \phi \rangle_A)(\varphi_B \langle \phi \rangle_B) \\ &\quad + 4(\langle \phi \rangle_A \varphi_A)(\varphi_B \varphi_B) + (\varphi_A \varphi_A)^2. \end{aligned} \quad (10.90)$$

To quadratic order, the action therefore takes the form

$$S = \int (dx) \left\{ \frac{1}{2}(\partial^\mu \varphi)(\partial_\mu \varphi) + \frac{1}{2} \varphi_A \left(m^2 + \frac{\lambda}{6} \langle \phi \rangle^2 \right) \varphi_A + \frac{\lambda}{6} \varphi_A (\langle \phi \rangle_A \langle \phi \rangle_B) \varphi_B + \dots \right\}. \quad (10.91)$$

If the action is evaluated at the minimum of the potential, substituting for the minimum $\langle\phi\rangle_A$, the quadratic masslike terms do not vanish, nor is any asymmetry created. The action is still invariant under rotations in A, B space, with a different mass matrix $\lambda/3\langle\phi\rangle_A\langle\phi\rangle_B$. However, if one postulates that it is favourable to select a particular combination for $\langle\phi\rangle_A$, e.g. let $A, B = 1, 2$ and $\langle\phi\rangle_1 = 0, \langle\phi\rangle_2 = \langle\phi\rangle$, thus breaking the symmetry between degenerate choices, then the quadratic terms become:

$$\frac{1}{2}\varphi_1\left(m^2 + \frac{\lambda}{6}\langle\phi\rangle^2\right)\varphi_1 + \frac{1}{2}\varphi_2\left(m^2 + \frac{\lambda}{2}\langle\phi\rangle^2\right)\varphi_2. \quad (10.92)$$

The first of these terms, evaluated at the minimum, vanishes, meaning that φ_1 is a massless excitation at the equilibrium solution. It is a Nambu–Goldstone boson, which results from the selection of a special direction. The rotational A, B symmetry of the fluctuating field φ_A is still present, but the direction of the average field is now chosen at all points.

In this two-dimensional rotational example, the special direction was chosen by hand, using the *ad hoc* assumption that the scalar field would have an energetically favoured ordered state. Clearly, one could have chosen any direction (linear combination of ϕ_A from the rotational invariance), and the result would be the same, due to the original symmetry. Since these are all equivalent, it takes only the energetic selection of any one of them to lead to an ordering, and thus spontaneous symmetry breaking. In the parametrization

$$\Phi = \frac{1}{\sqrt{2}}\rho e^{i\theta} \quad (10.93)$$

the symmetry properties of the action become even more transparent. The action is now:

$$S = \int (dx) \left[\frac{1}{2}(\partial^\mu\rho)(\partial_\mu\rho) + \frac{1}{2}m^2\rho^2 + \frac{\lambda}{4!}\rho^4 \right]. \quad (10.94)$$

This, assuming a stable average state $\rho \rightarrow \langle\rho\rangle + \rho$, gives, to quadratic order:

$$S = \int (dx) \left\{ \rho \left[-\square + m^2 + \frac{\lambda}{2}\langle\rho\rangle^2 \right] \rho + \langle\rho\rangle^2 \theta(-\square)\theta + \dots \right\} \quad (10.95)$$

The radial θ excitation is clearly massless. This parametrization has presented several technical challenges in the quantum theory however, so we shall not pursue it in detail.

The foregoing argument can be generalized to any continuous global group, either Abelian or non-Abelian. Suppose that the action

$$S = \int (dx) \{ T(\partial_\mu\phi_A) - V(\phi_A) \} \quad (10.96)$$

is invariant under a symmetry group G , of dimension d_G ; then, if it is energetically favourable for the field to develop a stable average $\langle\phi\rangle_A$ with restricted behaviour, such that

$$\phi_i \rightarrow \langle\phi\rangle_i + \varphi_i \tag{10.97}$$

for a sub-set of the components $i \in A$, there must a minimum in the potential, such that

$$\left. \frac{\partial V}{\partial \phi_i} \right|_{\phi=\langle\phi\rangle} = 0. \tag{10.98}$$

The field there splits into two parts:

$$\phi_A \rightarrow \begin{cases} \langle\phi\rangle_i + \varphi_i & \in H \\ \phi_i & \in G/H \end{cases} . \tag{10.99}$$

The first part has a stable average and small fluctuations around this value. The remainder of the components are unconstrained fluctuations, which are orthogonal in the group theoretical sense from the others. For the components with non-zero averages, one may expand the potential around the minimum:

$$V(\phi_A) = V(\phi_A) \Big|_{\phi_A=\langle\phi\rangle_A} + \left. \frac{\partial^2 V}{\partial \varphi_A \partial \varphi_B} \right|_{\phi_A=\langle\phi\rangle_A} \varphi_A \varphi_B + \dots \tag{10.100}$$

The form and value of the potential are unchanged by a group transformation G , since the action is invariant under G . Moreover, by assumption of a minimum, one must have

$$M_{AB} = \left. \frac{\partial^2 V}{\partial \varphi_i \partial \varphi_j} \right|_{\phi_A=\langle\phi\rangle_A} \geq 0. \tag{10.101}$$

To determine whether any of the components of this have to be zero, one uses the assumption that the average state is invariant under the sub-group H . Invariance under H means that

$$V(U_H \langle\phi\rangle) = V(\langle\phi\rangle) + \frac{\partial^2 V(\langle\phi\rangle)}{\partial \varphi_i \partial \varphi_j} \delta_H \langle\phi\rangle_i \delta_H \langle\phi\rangle_j + \dots; \tag{10.102}$$

thus, $\delta_H \langle\phi\rangle_i = 0$ and M_{ij}^2 is arbitrary, since the transformation itself is null-potent at $\langle\phi\rangle$. However, if one transforms the average state by an element which does not belong to the restricted group H , then $\delta_G \langle\phi\rangle \neq 0$, and

$$V(U_G \langle\phi\rangle) = V(\langle\phi\rangle) + \frac{\partial^2 V(\langle\phi\rangle)}{\partial \varphi_A \partial \varphi_B} \delta_G \langle\phi\rangle_A \delta_G \langle\phi\rangle_B + \dots \tag{10.103}$$

Thus, for any A, B which do not belong to i, j , the mass terms $M_{AB}^2 = 0$ for invariance of the potential. These are the massless modes. There are clearly $\dim G/H = d_G - d_H$ of these massless elements, which correspond to all of the fluctuations which are not constrained by the average state.

This argument does not depend on whether the group is Abelian or non-Abelian (except that the coset dimension G/H does not apply to groups like $U(1)$), only on the fact that a stable average emerges, picking out a special direction in group space. Since even a single group generator, corresponding to a single component of the field, generates a sub-group, the average field lies in a group of its own (the factor group). If the group H is an Abelian sub-group, such as Z_N , (generated by the Cartan sub-algebra of the full Lie algebra), then the resulting factor group shares the same algebra as the full group, only the centre of the group is broken. This changes the dimension of the representation, but does not change the universal cover group for the symmetry. If H is not an Abelian sub-group, then the basic algebra of the symmetry must also change.

The Nambu–Goldstone mechanism is a relative suppression of certain fluctuations, rather than a breakdown of fundamental symmetry. For example, in a crystal, with an R^n symmetry, the crystal lattice breaks up translations into R^n/Z_N , leading to massless vector fields, which are phonons.

It is not clear from the above that the choice of symmetry breaking potential is actually feasible: it has not been shown that the fluctuations around the average state are small enough to sustain the average value that was assumed. This requires a more lengthy calculation, using the generating functionals of chapter 6. Moreover, unless the result of the calculation can be determined entirely by quadratic terms, one is forced to use quantum field theory to calculate the expectation values, since there are questions of negative energies and probabilities which are only resolved by operator ordering in the second quantization. General theorems exist which prohibit the existence of Goldstone bosons, due to infra-red divergences, and thus global symmetry breaking in less than three spatial dimensions cannot occur by this mechanism [27, 97].

The occurrence of *spontaneous* symmetry breaking assumes that it will be possible to find a system in which the effective mass squared in the action is less than zero. Clearly no such fundamental fields exist: they would be tachyonic. However, composite systems, or systems influenced by external forces, can have effective mass-squared terms which have this property. This is exploited in heuristic studies of phase transitions, where one often writes the mass term as:

$$m^2(T) = \left(\frac{T - T_c}{T_c} \right) m_0^2 \quad (10.104)$$

which gives rise to a second-order phase transition at critical temperature T_c ($n > 2$), i.e. a change from an ordered average state at low temperature to a disordered state above the critical temperature.

10.8 Local symmetry breaking³

The predominance of gauge theories in real physical models leads one to ask whether symmetry breaking phenomena could occur in local gauge theories. Here one finds a subtly different mechanism, originally pointed out by Anderson [3], inspired by an observation of Schwinger [117], and rediscovered in the context of non-Abelian field theory by Higgs [68, 69, 70]. It is called the Anderson–Higgs mechanism, or simply the Higgs mechanism.

The action for this model is that of a complex scalar field coupled to the electromagnetic field. It is sometimes used as a simple Landau–Ginsburg model of super-conductivity (see section 12.6). It is also referred to as scalar electrodynamics. A straightforward non-Abelian generalization is used in connection with the Standard Model; this is discussed in many other references [136]. The action in complex form is written

$$S = \int (dx) \left\{ (D^\mu \Phi)^\dagger (D_\mu \Phi) + m^2 \Phi^\dagger \Phi + \frac{\lambda}{3!} (\Phi^\dagger \Phi)^2 + \dots + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right\}. \quad (10.105)$$

Here we have only written a Φ^4 interaction explicitly, with coupling constant λ . Other interactions are also possible depending on the criteria for the model. In the quantum theory, restrictions about renormalizability exclude higher powers of the field in $3 + 1$ dimensions. In $2 + 1$ dimensions one may add a term $\frac{8g}{6!} (\Phi^\dagger \Phi)^3$. Odd powers of the fields are precluded by the fact that the action must be real. The covariant derivative is usually written $D_\mu = \partial_\mu + ieA_\mu$. The conserved current generated by the gauge field A_μ is therefore

$$\frac{\delta S_\phi}{\delta A^\mu} = J_\mu = ie(\Phi^\dagger (D_\mu \Phi) - (D_\mu \Phi)^\dagger \Phi). \quad (10.106)$$

The action clearly has a basic $U(1)$ symmetry. An alternative form of the action is obtained by re-writing the complex field in terms of two real component fields ϕ_A , where $A = 1, 2$, as follows:

$$\Phi(x) = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2). \quad (10.107)$$

The covariant derivative acting on the fields can then be expanded in real and imaginary parts to give

$$D_\mu \phi_A = \partial_\mu \phi_A - e\epsilon_{AB} \phi_B A_\mu. \quad (10.108)$$

³ $\hbar = c = \mu_0 = \epsilon_0 = 1$ in this section.

The action then takes the more complicated form

$$S = \int (dx) \left\{ \frac{1}{2} (\partial^\mu \phi_A) (\partial_\mu \phi_A) - e (\partial^\mu \phi_A) \epsilon_{AB} A_\mu \phi_B + \frac{1}{2} e^2 \epsilon_{AB} \epsilon_{AC} \phi_B \phi_C A^\mu A_\mu + \frac{\lambda}{4!} (\phi_A \phi_A)^2 + \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right\}. \tag{10.109}$$

Expressed in this language, the conserved current becomes

$$J_\mu = e \epsilon_{AB} (\phi_A D_\mu \phi_B). \tag{10.110}$$

This shows the anti-symmetry of the current with respect to the field components in this $O(2)$ formulation.

Suppose, as before, that one component of the scalar field develops a constant non-zero expectation value $\phi_1 \rightarrow \langle \phi \rangle + \varphi_1$; the action can be expanded around this solution. Once again, this must be justified by an energy calculation to show that such a configuration is energetically favourable; is non-trivial and will not be discussed here. It is interesting to compare what happens in the presence of the Maxwell field with the case in the previous section. The part of the action, which is quadratic in $\varphi_1, \varphi_2, A_\mu$ is the dynamical part of the fluctuations. It is given by

$$S^{(2)} = \int (dx) \left\{ \frac{1}{2} \varphi_1 \left[-\square + m^2 + \frac{\lambda}{2} \langle \phi \rangle^2 \right] \varphi_1 + \frac{1}{2} \varphi_2 \left[-\square + m^2 + \frac{\lambda}{6} \langle \phi \rangle^2 + e^2 \langle \phi \rangle^2 \right] \varphi_2 + 2e\varphi_1 A^\mu (\partial_\mu \langle \phi \rangle) + \frac{1}{2} A_\mu \left[-\square + e^2 \langle \phi \rangle^2 \right] A^\mu \right\} \tag{10.111}$$

This may be diagonalized with the help of the procedure analogous to eqn. (A.11) in Appendix A. The identity

$$\frac{1}{2} \phi_2 A \phi_2 + B \phi_2 = \frac{1}{2} (\phi + BA^{-1}) A (\phi_2 + A^{-1} B) - \frac{1}{2} BA^{-1} B \tag{10.112}$$

with

$$A = \left[-\square + m^2 + \frac{\lambda}{6} \langle \phi \rangle^2 + e^2 \langle \phi \rangle^2 \right] \\ B = -2e\varphi_1 A^\mu (\partial_\mu \langle \phi \rangle) \tag{10.113}$$

results in an action of the form

$$S^{(2)} = \int (dx) \left\{ \frac{1}{2} \varphi_1 \left[-\square + m^2 + \frac{\lambda}{2} \langle \phi \rangle^2 \right] \varphi_1 + \frac{1}{2} A_\mu \left[(-\square + e^2 \langle \phi \rangle^2) g^{\mu\nu} + G \partial^\mu \partial^\nu \right] A_\nu \right\} \quad (10.114)$$

where G is a gauge-dependent term. The details of this action are less interesting than its general characteristics. Unlike the case of the global symmetry, there is only one remaining scalar field component. The component which corresponds to the Goldstone boson, disappears in the variable transformations and re-appears as a mass for the vector field. The lack of a Goldstone boson is also interesting, since it circumvents the problems associated with Goldstone bosons in lower dimensions $n < 3$ [27, 97]. Although it is only an idealized effective theory, this local symmetry breaking mechanism indicates that symmetry breaking is indeed possible when one relaxes the rigidity of a global group.

The transmutation of the massless scalar excitation into a mass for the vector field can be seen even more transparently in the *unitary gauge*. The unitary gauge is effected by the parametrization

$$\Phi = \frac{1}{\sqrt{2}} \rho e^{i\theta} \quad (10.115)$$

$$B_\mu = A_\mu + \frac{1}{e} \partial_\mu \theta \quad (10.116)$$

so that the action becomes

$$S = \int dV \left\{ \frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} (\partial^\mu \rho)(\partial_\mu \rho) + \frac{1}{2} e^2 \rho^2 B^\mu B_\mu + \frac{1}{2} m^2 \rho^2 + \frac{\lambda}{4!} \rho^4 \right\} \quad (10.117)$$

What looks like a gauge transformation by a phase θ is now a dynamical absorption of the Goldstone boson. This is sometimes stated by saying that the Goldstone boson is ‘eaten up’ by the gauge field, as if the photon were some elementary particular Pacman. A more field theoretical description is to say that the Goldstone mode modulates the fluctuations of the electromagnetic field, making them move in a wavefront. This wavefront impedes the fluctuations by an amount that depends upon the gauge coupling constant e . The result is an effective mass for the gauge fluctuations, or a gap in their spectrum of excitations. However one states it, the Goldstone field ceases to be a separate

excitation due to the coupling: its modulation of the vector field's zero point energy breaks the gauge invariance of the fluctuations and it re-appears, with a new status, as the extra mode of the vector field.

It cannot be emphasized enough that the assumption that there exists a stable average state of lower symmetry than the fluctuations of the theory is *ad hoc*, and its consistency has to be proven. Even today, this remains one of the toughest challenges for quantum field theory.

10.9 Dynamical symmetry breaking mechanisms

The Nambu–Goldstone or Anderson–Higgs models of symmetry breaking cannot be fundamental theories, because they do not explain how the mass-squared terms, in their Lagrangians, can become negative. As such, they must be regarded as effective actions for deeper theories. Moreover, their apparent reliance on the existence of an arbitrary scalar field has been controversial, since, in spite of the best efforts of particle physicists, no one has to date observed a Higgs scalar particle. The introduction of a scalar field is not the only way in which gauge symmetries can be broken, however. At least two other possibilities exist. Both rely on quantum dynamical calculations, but can be mentioned here.

One such mechanism was suggested in connection with field theories on topologically non-trivial spacetimes (e.g. the torus), based on an idea by Ford [52], that non-trivial average states, such as vortices could occur around topological singularities in spacetime. The main idea is that a gauge field $A_\mu \rightarrow \langle A_\mu \rangle + A_\mu$ (either Abelian or non-Abelian) can acquire a non-zero expectation value around a hole in spacetime. In simply connected spacetimes (without holes), such constant vector field configurations are gauge-equivalent to zero and thus have no invariant meaning. However, around a topological singularity, such transformations are restricted by the cohomology of the manifold. One example is that of a periodic crystal, which has the same boundary conditions as the surface of a torus, and is therefore relevant in solid state physics.

In the Abelian theory, the phenomenon is a purely classical, statistical effect, though for non-Abelian symmetries the non-linearity makes it the domain of quantum field theory. It is equivalent to there being a constant magnetic flux through the centre of the hole. In some theories, such expectation values might occur spontaneously, by the dynamics of the model (without having to assume a negative mass squared *ad hoc*). In the Abelian case, this results only in a phase. However, it was later explored in the context of non-Abelian symmetries by Hosotani [72] and Toms [129] and developed further in refs. [17, 19, 20, 21, 32, 33]. Such models are of particular interest in connection with grand unified theories, such as Kaluza–Klein and string theory, where extra dimensions are involved. Topological singularities also occur in lower dimensions in the form of vortices and the Aharonov–Bohm effect.

The second mechanism is the Coleman and Weinberg mechanism [28], which is a purely quantum effect for massless fields, whereby a non-trivial average state can be created truly spontaneously, by the non-linearities of massless scalar electrodynamics. Quantum fluctuations themselves lead to the attainment of an ordered state. It is believed that this mechanism leads to a first-order phase transition [66, 86], rather than the second-order transitions from the Goldstone and Higgs models.