2 Broken Symmetries

Global symmetries have two important roles to play in physics. First, they lead to conservation laws through Noether's theorem. Second, if the symmetry is non-Abelian then it leads to a degeneracy in the spectrum, as the states of the theory necessarily furnish a representation of the symmetry. This is familiar from the quantum treatment of the hydrogen atom where states sit in multiplets of the $SO(3)$ rotation group of dimension $2l + 1$ where l is the angular momentum.

But there are other ways in which symmetries can affect the dynamics of a theory. And this happens when symmetries are "broken".

There are actually two different meanings to the phrase "broken symmetry", both of which arise in the context of the Standard Model. The first, sometimes called *explicit breaking*, is when there are terms in the action that are not invariant under the symmetry. Strictly speaking, this is the same as not having a symmetry at all. But the symmetry can still be a useful fiction if the terms that break it are, in some sense, small so that we have an approximate symmetry. In this case, it might be that some quantity is almost conserved, meaning that violations of the conservation law happen rarely. Or it could be that the degenerate multiplets that arose when the symmetry was exact are split by some small amount. This happens, for example, if we place the hydrogen atom in a magnetic field so that the rotation symmetry is broken. Then the $2l+1$ states which were previously all degenerate get slightly split by the Zeeman effect.

In the Standard Model, we will see several examples of approximate symmetries, including isospin and its extension to an $SU(3)$ flavour symmetry known as the *eightfold way*, as well as chiral symmetry. Both of these will be explained in section 3.

The second meaning of the term "broken symmetry" refers to a more subtle and, ultimately, more powerful phenomenon. This arises when the theory is invariant under a symmetry, but the ground state is not. This situation is referred to as *spontaneous symmetry breaking*. The purpose of this section is to explain when this happens and what the consequences are.

Spontaneous symmetry breaking is one of those lovely ideas that crosses into many different areas of physics. It was one of the major themes of the lectures on Statistical Field Theory where it underlies Landau's theory of phase transitions. It also arises in many places in condensed matter physics, from magnets to superconductors. For example, sound waves in a solid can be viewed as the consequence of spontaneous breaking of translation symmetry by the underlying lattice. Spontaneous symmetry breaking also occurs in at least two different contexts in the Standard Model.

2.1 Discrete Symmetries

The idea of spontaneous symmetry breaking is not something new: it appears in some simple classical mechanics systems.

Consider a real, classical degree of freedom $\phi(t)$ with action given by

$$S = \int dt \, \left(\frac{1}{2}\dot{\phi}^2 - V(\phi)\right) \quad \text{with} \quad V(\phi) = \frac{m^2}{2}\phi^2 + \frac{\lambda}{4}\phi^4 \ . \tag{2.1}$$

In Newtonian mechanics, we would think of $\phi(t)$ as the position of a particle and usually denote it as $x(t)$. We're going to avoid calling the degree of freedom x because we'll soon make the leap to field theory where x becomes an argument of the field, $\phi(\mathbf{x}, t)$. But you should feel free to think of $\phi(t)$ as the position of a particle.

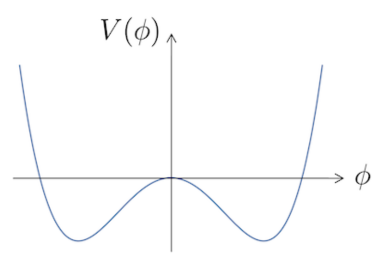
The potential (2.1) enjoys a discrete $\mathbb{Z}_2$ symmetry under which

$$\mathbb{Z}_2 : \phi \mapsto -\phi \ . \tag{2.2}$$

In classical mechanics, where ϕ is the position of the particle, this symmetry is called "parity" but we'll avoid this name because, again, in the context of field theory parity acts differently (as we saw in Section 1.4).

The issue of spontaneous symmetry breaking is all about the sign of the first term in the potential. When $m^2 > 0$, the potential has a minimum at $\phi = 0$. This is the one point that is invariant under the symmetry $\phi \mapsto -\phi$ and we say that the symmetry is unbroken.

In contrast, if $m^2 < 0$ then the ϕ^2 term in (2.1) comes with a negative coefficient and the point $\phi = 0$ is now a local maximum rather than a minimum, as shown in the figure. This is the *double well potential*. The minimum lies at



$$\phi = \pm v \equiv \pm\sqrt{-\frac{m^2}{\lambda}} \ . \tag{2.3}$$

We see that two, related things occur. First, there is not a unique ground state: there are two. Second, neither ground state is invariant under the $\mathbb{Z}_2$ symmetry (2.2). Instead, the symmetry exchanges the two ground states. This is our first, admittedly somewhat trivial, example of spontaneous symmetry breaking. But there is an important lesson that will carry over to more complicated situations: if a discrete symmetry is spontaneously broken, then the theory has multiple, ground states with a potential barrier between them. Acting with the symmetry then transforms us among the ground states.

Suppose that you sit in one of the two ground states, and look only at small oscillations about the minimum. What do you see? We write the potential (2.1) as

$$V(\phi) = \frac{\lambda}{4}(\phi^2 - v^2)^2 + \text{constant}\ . \tag{2.4}$$

We take ourselves to sit near the ground state $\phi = +v$ and expand

$$\phi(t) = v + \sigma(t)\ . \tag{2.5}$$

We can then substitute this back into the potential (2.4) to get

$$V(\sigma) = \frac{\lambda}{4}(2v\sigma + \sigma^2)^2 = \lambda\left(v^2\sigma^2 + v\sigma^3 + \frac{\sigma^4}{4}\right) + \text{constant}\ . \tag{2.6}$$

We see that, while the full potential $V(\phi)$ has the $\mathbb{Z}_2$ symmetry, if you're trapped near one of the minima then you know nothing about it. The action for small oscillations includes the σ^3 term and most certainly isn't invariant under $\sigma \mapsto -\sigma$. This is the sense in which the $\mathbb{Z}_2$ symmetry is hidden, or broken, about any given ground state. The consequence of the symmetry, when broken, is only to generate multiple ground states.

2.1.1 Quantum Tunnelling

The discussion above is straightforward enough and holds for classical particle mechanics. But quantum mechanics brings an extra twist. This is because there is no spontaneous symmetry breaking in quantum mechanics! The ground state is always invariant under the $\mathbb{Z}_2$ symmetry. In fact, all energy eigenstates are invariant under the $\mathbb{Z}_2$ symmetry.

You might be tempted to construct a ground state that is localised near one or other of the minima, say a wavefunction of the form

$$\psi_{\text{left}}(\phi) \approx \exp\left(-\frac{\sqrt{\lambda}v}{2}(\phi + v)^2\right) \quad \text{or} \quad \psi_{\text{right}}(\phi) \approx \exp\left(-\frac{\sqrt{\lambda}v}{2}(\phi - v)^2\right). \tag{2.7}$$

But neither of these are eigenstates of the $\mathbb{Z}_2$ symmetry, and neither are eigenstates of the Hamiltonian. Indeed, if you were to place the system in, say, $\psi_{\text{left}}(\phi)$ then the wavefunction will leak through the barrier in a process known as *quantum tunnelling*.

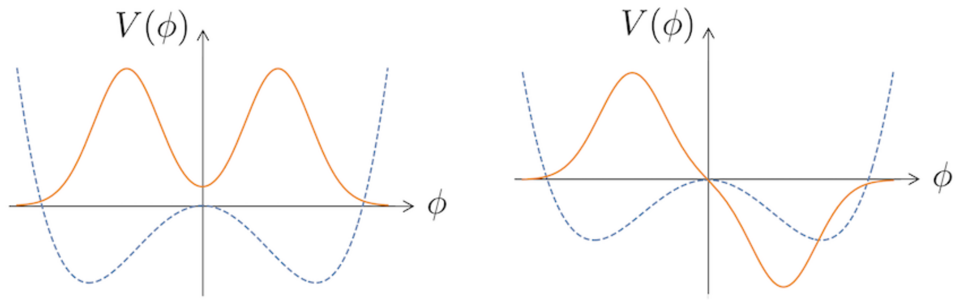


Figure 2. On the left: the ground state of the double well potential. On the right: the first excited state.

Instead, the true ground state wavefunction takes the approximate form

$$\psi_{\text{ground}}(\phi) \approx \psi_{\text{left}}(\phi) + \psi_{\text{right}}(\phi) \ . \tag{2.8}$$

The ground state has no zeros other than at $\phi \to \pm\infty$. Meanwhile, the first excited state is

$$\psi_{\text{excited}}(\phi) \approx \psi_{\text{left}}(\phi) - \psi_{\text{right}}(\phi) \ . \tag{2.9}$$

This has a single node, meaning that it crosses the axis once. The n^{th} excited state has n nodes. (See the lectures on Quantum Mechanics for more discussion of these facts.) The ground state and first excited state are shown in Figure 2.

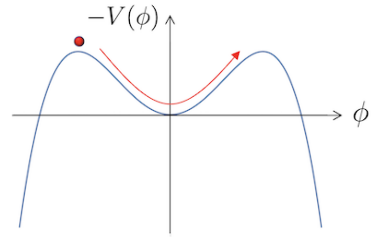
There is another way to see tunnelling that will prove useful when we turn to quantum field theory shortly. We want to compute the amplitude for a particle to start in one minimum, say $\phi = -v$, and end up at the other minimum $\phi = +v$. We can do this using the path integral. After Wick rotating to work with imaginary time $\tau = it$, we have

$$\langle +v|e^{-H\tau}|-v\rangle = \int \mathcal{D}\phi \ e^{-S_E[\phi]} \ . \tag{2.10}$$

Here $S_E[\phi]$ is the "Euclidean action", meaning that is differs from (2.1) by a minus sign.

$$S_E[\phi] = \int d\tau \ \left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right) \ . \tag{2.11}$$

To compute the amplitude (2.10), we should evaluate the path integral on paths that start in the left-hand vacuum and end up at the right-hand vacuum. We can get some intuition for this by noting that the Euclidean action (2.11) simply flips the sign of the potential term, so if we wished to view it as a classical mechanics system then it describes a particle rolling in the inverted potential $-V(\phi)$. We're then looking for paths that start perched on the left-hand peak, roll down to the minimum, and then rise again to end on the right-hand peak, as shown in the figure.

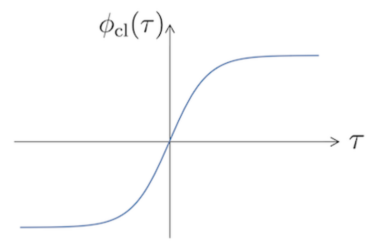


The path integral instructs us to integrate over all such paths. But, in the saddle point approximation, we expect the dominant contribution to come from paths that obey the classical equation of motion,

$$\ddot{\phi} = \lambda\phi(\phi^2 - v^2)\ . \tag{2.12}$$

This equation has a rather nice analytic solution that does what we want, namely

$$\phi_{\rm cl}(\tau) = v\tanh\left(\sqrt{\frac{\lambda v^2}{2}}\tau\right)\ . \tag{2.13}$$



The profile is shown in the figure to the right. It interpolates from $\phi = -v$ to $\phi = +v$, with the interesting stuff happening over a time period $\Delta\tau \sim 1/\sqrt{\lambda v^2} \sim 1/|m|$. We can evaluate the Euclidean action (2.11) on this solution to get

$$\begin{aligned} S_{\rm cl} &= \int_{-\infty}^{+\infty} d\tau\ \left(\frac{1}{2}\dot{\phi}_{\rm cl}^2 + \lambda(\phi_{\rm cl}^2 - v^2)^2\right) \\ &= \frac{\lambda v^4}{2}\int_{-\infty}^{+\infty} d\tau\ \frac{1}{\cosh^4(\sqrt{\lambda v^2/2}\tau)} \\ &= \frac{2}{3}\sqrt{2\lambda}v^3\ . \end{aligned} \tag{2.14}$$

This can be viewed as a measure of how difficult it is to tunnel under the barrier. As the barrier gets bigger (so λ increases) or the minima get further apart (so v^2 increases), the classical action $S_{\rm cl}$ also increases. This then gives our first guess at the amplitude to tunnel from one minimum to the other,

$$\lim_{\tau\to\infty}\langle +v|e^{-H\tau}|-v\rangle = Ke^{-S_{\rm cl}}. \tag{2.15}$$

Here K is some overall constant that masks all manner of sins that we've swept under the rug. In fact, to do this calculation correctly, we should really be summing over trajectories that bounce back and forth many times. One then finds, in the limit of large T, that you have just as much chance of being in the vacuum $\phi = -v$ as you do of being in the vacuum $+v$. This is the statement that there is no spontaneous symmetry breaking in quantum mechanics. Moreover, you find that the energy difference between the ground state and first excited state is given by

$$E_{\text{excited}} - E_{\text{ground}} \approx \sqrt{\lambda v^2} e^{-S_{\text{cl}}} \ . \tag{2.16}$$

The splitting of the two states is exponentially suppressed.

With these ideas in mind, we can now return to what we really care about: quantum field theory.

2.1.2 Discrete Symmetry Breaking in Quantum Field Theory

We now extend our double well discussion to field theory. Now $\phi(x)$ is a function of spacetime. The action (2.1) is replaced by

$$S = \int d^4x \ \left(\frac{1}{2}\partial_\mu \phi \, \partial^\mu \phi - V(\phi)\right) \quad \text{with} \quad V(\phi) = \frac{m^2}{2}\phi^2 + \frac{\lambda}{4}\phi^4 \ . \tag{2.17}$$

Again, we have a $\mathbb{Z}_2$ symmetry $\phi \mapsto -\phi$ and, when $m^2 < 0$, we have a double well potential with two minima at $\phi = \pm v = \pm\sqrt{-m^2/\lambda}$. We want to ask: is this symmetry spontaneously broken or not?

Quantum field theory is an extension of quantum mechanics (the clue is in the name) so we might think that tunnelling would again mean that there is no spontaneous symmetry breaking. But that's not the way things work. This is one situation where field theory differs from quantum mechanics and our classical intuition is better. The quantum field theory really does have two ground states, in which the vacuum expectation value of the field is given by

$$\langle \phi \rangle = \pm v \ . \tag{2.18}$$

To see why quantum field theory is different from common or garden quantum mechanics, we can return to the tunnelling calculation that we saw above. We can again compute the amplitude to go from one putative ground state to another,

$$\langle +v | e^{-H\tau} | -v \rangle = \int \mathcal{D}\phi \ e^{-S_E[\phi]} \ . \tag{2.19}$$

The Euclidean action $S_E[\phi]$ is now

$$S_E[\phi] = \int d\tau d^3x \, \left(\frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi + V(\phi)\right) \; . \tag{2.20}$$

In the saddle point approximation, the amplitude is dominated by the classical solutions which obey

$$\partial^2\phi = \lambda\phi(\phi^2 - v^2) \; . \tag{2.21}$$

This is the same as (2.12), but with the $\ddot{\phi}$ term replaced by the Laplacian on (Euclidean) spacetime, $\partial^2 = \partial_\tau^2 + \nabla^2$. We still have the same solution as before,

$$\phi_{\rm cl}(\tau) = v \tanh\left(\sqrt{\frac{\lambda v^2}{2}}\tau\right) \; . \tag{2.22}$$

The field varies in (Euclidean) time τ but is constant in space. So far, everything runs in parallel to the quantum mechanics argument. But now we compute the classical action of this solution. It is

$$S = \int d\tau d^3x \, \left(\frac{1}{2}\partial_\mu\phi_{\rm cl}\,\partial^\mu\phi_{\rm cl} + V(\phi_{\rm cl})\right) = \mathcal{V} S_{\rm cl} \; . \tag{2.23}$$

Here $S_{\rm cl}$ is the quantum mechanical action (2.14) while $\mathcal{V}$ is the volume of space. But, if we're working in uncompactified Minkowski space then $\mathcal{V} = \infty$. This means that both the tunnelling amplitude (2.15) and the energy splitting of the ground states (2.16) are proportional to

$$e^{-\mathcal{V}S_{\rm cl}} \to 0 \quad \text{as} \quad \mathcal{V} \to \infty \; . \tag{2.24}$$

It's obvious what's going on here. In quantum field theory, the ground state of the field in one minimum is, say, $\phi(\mathbf{x}) = +v$ for all $\mathbf{x}$. If you want to tunnel to the other minimum, $\phi(\mathbf{x}) = -v$, then you have to shift the value of the field at every point in space. But that takes effort and quantum tunnelling is not up to the task. It costs an infinite amount of action and so does not occur.

This means that while discrete symmetries cannot be spontaneously broken in quantum mechanics, they can be broken in quantum field theory. The suppression is by the volume factor, so if we're working with quantum field theory on some compact space, rather than infinite volume Minkowski space, then tunnelling reappears. However, if the space is macroscopically large then the suppression factor $e^{-\mathcal{V}S_{\rm cl}}$ may be so tiny that, for all intents and purposes, we can think of the symmetry as broken.

The upshot of this argument is that the quantum field theory (2.17) in $d = 3 + 1$ dimensions (and, indeed, in any dimension greater than $d = 0 + 1$) has two ground states, $|+v\rangle$ and $|-v\rangle$, distinguished by the expectation value of $\phi(x)$ which acts as an *order parameter* to tell us which vacuum we live in,

$$\langle \pm v|\phi(x)| \pm v\rangle = \pm v \quad \text{and} \quad \langle \pm v|\phi(y)| \mp v\rangle = 0. \tag{2.25}$$

This is a story that generalises to other discrete symmetries. For example, if you find yourself with a quantum field theory with $\mathbb{Z}_N$ symmetry which is spontaneously broken, then you will have N ground states that will be permuted into each other by the action of the symmetry.

The Meaning of a Tachyon

Tachyons are mythological beasts in physics. When we first learn special relativity, certain unscrupulous teachers may tell you that a tachyon is a particle with $m^2 < 0$ which is forced forever to travel faster than the speed of light. This is, of course, nonsense.

In field theory, a tachyon is nothing mysterious. Our potential above has $m^2 < 0$ but there is certainly nothing flying around faster than light. Instead, it signals that the point $\phi = 0$ is a maximum of the potential, rather than a minimum. This is the true meaning of a tachyon in field theory: it is telling us that the chosen vacuum is unstable. It's our job to find a better, stable vacuum.

That's not hard in the example above. We just need to expand around one of the minima of the potential, rather than the maximum. In fact, we already did this calculation in (2.6). If we write $\phi(x) = v + \sigma(x)$, then we find a potential for σ given by

$$V(\sigma) = \lambda \left(v^2\sigma^2 + v\sigma^3 + \frac{1}{4}\sigma^4 \right) \ . \tag{2.26}$$

We can read off the mass of particles in the theory from the quadratic term. Any physical excitation has mass $M^2 = 2\lambda v^2$. The mass is real and positive and decidedly not exotic in any way.

Domain Walls

The presence of a spontaneously broken symmetry often implies the existence of some novel excitation in the theory. In the present case, this is a *domain wall*, a field configuration that interpolates from one vacuum to the other.

Indeed, we've already met the classical solution that does the job. We just need to repurpose the tunnelling solution (2.22) by replacing the imaginary time τ with one of the spatial coordinates $\mathbf{x} = (x, y, z)$. For example, the classical field configuration

$$\phi(z) = v \tanh\left(\sqrt{\frac{\lambda v^2}{2}} z\right) \tag{2.27}$$

solves the equations of motion of the original Lorentzian action (2.17). This solution interpolates from the vacuum $\phi = -v$ at $z \to -\infty$ to the vacuum $\phi = +v$ at $z \to +\infty$. It describes an excitation of the field, localised around $z = 0$, but extended in the x- and y-directions. This is the domain wall.

The domain wall has finite energy density $\mathcal{E}$ which, it is easy to see, coincides with the action $S_{\rm cl}$ of the same configuration in quantum mechanics. We computed this in (2.14) and found

$$\mathcal{E} = \frac{2}{3}\sqrt{2\lambda} v^3 \ . \tag{2.28}$$

Although the domain wall has finite energy density, it has infinite energy because it stretches to infinity in the (x, y)-plane. An exception to this statement is if we are considering domain walls in $d = 1 + 1$ dimensions where there is nowhere else for them to stretch. In this case the domain walls have finite energy and should be viewed as a kind of particle in the theory.

Back in $d = 3 + 1$ dimensions, we can straightforwardly consider variations of this classical configuration (2.27) in which the domain wall forms a sphere of radius R, containing one vacuum $\phi = -v$ inside, and the other vacuum $\phi = +v$ outside. This now has finite energy, given by $E = 4\pi R^2 \mathcal{E}$. However, such a static configuration will no longer solve the equation of motion because the domain wall has tension and will want to contract. To find the classical solution, we will have to solve the full time-dependent partial differential equation.

We can also get some sense for what happens to these configurations in the quantum theory. We can build a Fock space of states above either of the two ground states by exciting the field $\phi(x) = \pm v + \sigma(x)$. As we've noted, this creates particles of mass $M = \sqrt{2\lambda v^2}$. The Hilbert space of the theory decomposes as

$$\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_- \ . \tag{2.29}$$

This is *not* a tensor product, which would mean that we have to choose one state from $\mathcal{H}_+$ and another from $\mathcal{H}_-$ to specify the full state. Instead, it's a tensor sum: we must pick *either* a state from $\mathcal{H}_+$ *or* a state from $\mathcal{H}_-$. Those states $|\psi\rangle \in \mathcal{H}_+$ obey

$$\langle\psi|\phi(0,\mathbf{x})|\psi\rangle = +v \quad \text{for} \quad |\mathbf{x}| \to \infty \ . \tag{2.30}$$

This is telling us that we necessarily approach the vacuum $|+v\rangle$ when we're far away. However, this doesn't mean that the excitations about one ground state know nothing about the other ground state. By piling many ϕ excitations on top of each other, it's quite possible to carve out a region of one vacuum inside another, and have excited states $|\psi\rangle \in \mathcal{H}_+$ that obey, for example,

$$\langle\psi|\phi(0,\mathbf{x})|\psi\rangle = \begin{cases} -v & \text{for } |\mathbf{x}| < R \\ +v & \text{as } |\mathbf{x}| \to \infty \end{cases} \ . \tag{2.31}$$

These kind of states are what become of our classical, spherical domain wall.

Cluster Decomposition

We know that the field theory has two ground states $|\pm v\rangle$, but you might wonder why we're necessarily forced to work with these states. What's stopping us taking the linear combinations

$$|0_\pm\rangle = \frac{1}{\sqrt{2}}\Big(|+v\rangle \pm |-v\rangle\Big) \tag{2.32}$$

as our ground states? This is a superposition of a state in $\mathcal{H}_+$ and a state in $\mathcal{H}_-$.

In fact, $|0_\pm\rangle$ are not the right states to work with. There are two arguments for this. The first is a little handwavey. Suppose that we perturb our original Lagrangian by some term $\Delta\mathcal{L}$ that breaks the $\mathbb{Z}_2$ symmetry. This will mean that one of the states $|\pm v\rangle$ has lower energy and is the true ground state. In the limit that we send the coefficient of $\Delta\mathcal{L}$ to zero, we will remain in the ground state, either $|+v\rangle$ or $|-v\rangle$.

This argument seems more compelling for condensed matter systems, where you can well imagine that there are many different perturbations (say, background magnetic fields) that would break the $\mathbb{Z}_2$ symmetry. The argument is less convincing in the context of particle physics where it's not at all clear what these additional terms might be. (Some balm comes from a conjecture that, once we take gravity into account, there are no exact global symmetries so there must, in fact, be some irrelevant symmetry breaking term lurking in the wings.)

There is a second, more important argument for why the states $|0_\pm\rangle$ defined in (2.32) are not the right ground states. This is a property known as *cluster decomposition* which is a way of capturing the locality of field theory. If you sit in some vacuum state $|\text{vac}\rangle$ and compute the two-point function of two operators, $A(x)$ and $B(y)$ then, when x and y are spacelike separated, the expectation value should decompose into

$$\langle\text{vac}|A(x)B(y)|\text{vac}\rangle \to \langle\text{vac}|A(x)|\text{vac}\rangle\,\langle\text{vac}|B(y)|\text{vac}\rangle \quad \text{as} \quad |x-y|\to\infty. \tag{2.33}$$

Now, on general grounds you can argue that, when x and y are far separated, we must have

$$\langle\text{vac}|A(x)B(y)|\text{vac}\rangle \to \sum_n \langle\text{vac}|A(x)|n\rangle\,\langle n|B(y)|\text{vac}\rangle \tag{2.34}$$

where $|n\rangle$ run over all possible vacuum states. But for cluster decomposition to hold, we want this to project onto the specific vacuum state $|n\rangle = |\text{vac}\rangle$ that we started in.

We can check this criterion for our theory with spontaneous symmetry breaking and the choice $A = B = \phi$. If we pick the state $|+v\rangle$ then, using the fact that $\langle +v|\phi(x)|-v\rangle = 0$, we have

$$\langle +v|\phi(x)\phi(y)|+v\rangle \to \langle +v|\phi(x)|+v\rangle\,\langle +v|\phi(y)|+v\rangle = v^2 \tag{2.35}$$

So this indeed obeys cluster decomposition. In contrast, if we work in the state $|0_+\rangle$ defined in (2.32) then you can check that

$$\langle 0_+|\phi(x)|0_+\rangle = \langle 0_-|\phi(x)|0_-\rangle = 0 \quad \text{and} \quad \langle 0_+|\phi(x)|0_-\rangle = v\ . \tag{2.36}$$

We then have

$$\langle 0_+|\phi(x)\phi(y)|0_+\rangle \to \langle 0_+|\phi(x)|0_-\rangle\,\langle 0_-|\phi(y)|0_+\rangle = v^2 \tag{2.37}$$

This does not obey cluster decomposition because the vacuum $|0_-\rangle$ that we need to insert in the middle differs from the vacuum $|0_+\rangle$ that we started with.

2.2 Continuous Symmetries

The story of symmetry breaking is rather different, and more powerful, when the symmetry in question is a continuous symmetry. Here we start by giving a couple of examples before we describe the general result known as Goldstone's theorem.

We'll work in quantum field theory. As in the previous section, there is some tension between spontaneous symmetry breaking in quantum field theory and what we know about the behaviour of wavefunctions in quantum mechanics, but we'll put this on hold for now and return to it in Section 2.2.4.

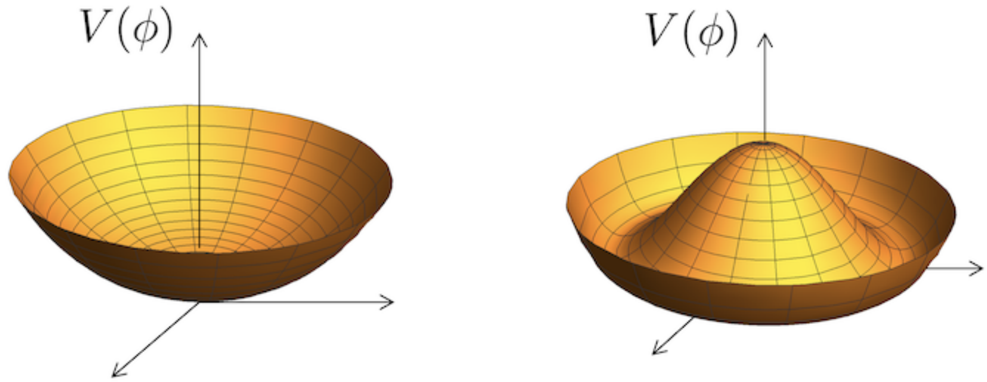


Figure 3. On the left: the potential with $m^2 > 0$. On the right, the Mexican hat potential with $m^2 < 0$.

To start, consider a complex scalar field $\phi(x)$ in $d = 3 + 1$ dimensions with action

$$S = \int d^4x \, \left(\partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi, \phi^\dagger) \right) \quad \text{with} \quad V(\phi, \phi^\dagger) = m^2 |\phi|^2 + \frac{1}{2} \lambda |\phi|^4 \ . \tag{2.38}$$

The action is constructed so that it a enjoys $U(1)$ global symmetry which rotates the phase of ϕ,

$$\phi(x) \to e^{i\alpha} \phi(x) \ . \tag{2.39}$$

Again, the physics depends on the sign of the m^2 term in the potential. The two different cases, with $m^2 > 0$ and $m^2 < 0$ are shown in Figure 3. In the former case, there is little interesting to say: you expand around the vacuum $\phi = 0$ and, after quantisation, find interacting particles of mass m with the $U(1)$ symmetry implying the usual conservation law. Here our interest is in the case $m^2 < 0$.

The potential with $m^2 < 0$ is sometimes called the "Mexican hat potential" because, you know, . It also looks like the bottom of a wine bottle. The defining feature is that there are not isolated minima, but instead an infinite number of ground states, defined by

$$|\phi|^2 = -\frac{m^2}{\lambda} \ . \tag{2.40}$$

We define the *vacuum manifold* $\mathcal{M}_0$ to be the space of field configurations which have minimum energy. For the double well potential of Section 2.1, the vacuum manifold was just two points. Now, the vacuum manifold is the set of solutions to (2.40) which is a circle,

$$\mathcal{M}_0 = \mathbf{S}^1 \ . \tag{2.41}$$

To see what this buys us, we can write the complex field in polar coordinates, with

$$\phi(x) = r(x)e^{i\theta(x)} \ . \tag{2.42}$$

This is a slightly dangerous thing in quantum field theory, where we usually assume that fields can take any value. In writing (2.42), we need to remember that $r(x) \geq 0$ and $\theta(x) = \theta(x) + 2\pi$. Nonetheless, we can proceed for now and keep this in the back of our minds.

Substituting the polar decomposition into the original action (2.38), we have

$$S = \int d^4x \ \left(\partial_\mu r \partial^\mu r + r^2 \partial_\mu \theta \partial^\mu \theta - \frac{\lambda}{2}(r^2 - v^2)^2\right) \tag{2.43}$$

where, as in the last section, we've introduced $v^2 = -m^2/\lambda$. Now we can read off the physics. The ground state of the system sits at $r(x) = +v$. If we expand about this vacuum by writing $r(x) = v + \sigma(x)$ then the action becomes

$$S = \int d^4x \ \left(\partial_\mu \sigma \partial^\mu \sigma + (v+\sigma)^2 \partial_\mu \theta \partial^\mu \theta - \frac{\lambda}{2}\sigma^2(\sigma + 2v)^2\right) \tag{2.44}$$

From this, we can read off the physics. In particular, the $\sigma(x)$ excitations have mass $M^2 = 2\lambda v^2$. These are radial oscillations of the field, that go back and forth in the potential.

To pick a vacuum, we also need to specify a value for the angular scalar field $\theta(x)$. But there is no preferred choice here. Once we've set $r(x) = v$, the different constant values of $\theta(x)$ parameterise the vacuum manifold $\mathcal{M}_0 = \mathbf{S}^1$. If this was quantum mechanics, then the wavefunction would simply spread over the $\mathbf{S}^1$. But things are different in quantum field theory, a fact that we will discuss further in Section 2.2.4, and each point on $\mathcal{M}_0$ corresponds to a different ground state of the theory. To specify the ground state, we have to pick one such point. It doesn't matter which point we pick because the physics will be the same in each. But, nonetheless, we have to pick one.

Whatever choice of ground state we make, say $\theta(x) = 0$, will spontaneously break the $U(1)$ symmetry (2.39) which acts as

$$\theta(x) \rightarrow \theta(x) + \alpha \ . \tag{2.45}$$

In fact, we see that the symmetry acts by taking us from one point on $\mathcal{M}_0$ to another.

Finally, we can look at the dynamics of the field $\theta(x)$ that parameterises $\mathcal{M}_0$. From the action (2.43), we see that there is no potential term for θ, a fact which simply follows from the $U(1)$ invariance of the potential. If we ignore the coupling to σ, then the θ field is governed by the simple Lagrangian

$$\mathcal{L} = v^2 \partial_\mu \theta \, \partial^\mu \theta \ . \tag{2.46}$$

This is a Lagrangian for a massless scalar field, albeit one that is slightly unusual because θ is a periodic variable. The existence of this massless scalar field is a direct consequence of the spontaneous breaking of the $U(1)$ global symmetry. As we will see, this is a general story: whenever a continuous global symmetry is spontaneously broken, there will be massless scalar fields. These fields are called *Goldstone bosons*.

Goldstone bosons can't have potential terms: only derivative terms. But that's not to say that they're totally boring. There can still be interactions, both among themselves (as we will see in later examples) and with other fields. For example, if we expand out $r(x) = v + \sigma(x)$ in (2.43) then we see that there are interaction terms between the massive scalar σ and the massless Goldstone boson θ that take the form $\sigma(\partial\theta)^2$ and $\sigma^2(\partial\theta)^2$. This means that a σ particle can decay to two Goldstone modes. However, if we look at energies $E \ll \sqrt{\lambda v^2}$, which is the mass of the σ particle, then the only field in town is the massless Goldstone mode, whose dynamics is governed by (2.46).

2.2.1 The $O(N)$ Sigma Model

Here's a generalisation of the ideas above. We take a collection of N real scalar fields $\phi^a(x)$, with $a = 1, \ldots, N$, and consider the following action

$$S = \int d^4x \, \left(\frac{1}{2}\partial_\mu \phi^a \partial^\mu \phi^a - V(\phi)\right) \quad \text{with} \quad V(\phi) = \frac{1}{2}m^2 \phi^a \phi^a + \frac{1}{4}\lambda(\phi^a\phi^a)^2 \ . \tag{2.47}$$

This action is constructed to have an $O(N)$ symmetry, under which the ϕ^a rotate. For $N = 2$, it coincides with the action (2.38) for a complex scalar field whose real and imaginary parts are ϕ^1 and ϕ^2.

Spontaneous symmetry breaking occurs when $m^2 < 0$ and the potential again looks like a Mexican hat but for someone with a higher dimensional head. The minima of the potential obey

$$\phi^a \phi^a = v^2 := -\frac{m^2}{\lambda} \ . \tag{2.48}$$

This is simply the equation for an $(N-1)$-dimensional sphere, and defines the vacuum manifold of the theory

$$\mathcal{M}_0 = \mathbf{S}^{N-1} \ . \tag{2.49}$$

The vacuum of the theory is one point on $\mathcal{M}_0$. It doesn't matter which one. Suppose that we pick the "south pole", so that the vacuum is $\phi^a = (0, 0, \ldots, 0, v)$. Now we can look at fluctuations around this vacuum by writing

$$\phi^a(x) = \big(\pi^1(x), \ldots, \pi^{N-1}(x), v + \sigma(x)\big) \ . \tag{2.50}$$

If we substitute this into the action (2.47), we find

$$S = \int d^4x \ \left(\frac{1}{2}\partial_\mu \pi^a \partial^\mu \pi^a + \frac{1}{2}\partial_\mu \sigma \, \partial^\mu \sigma - V(\pi^a, \sigma)\right) \tag{2.51}$$

with

$$V(\pi^a, \sigma) = \lambda v^2 \sigma^2 + \lambda v \sigma \left(\sigma^2 + \pi^a \pi^a\right) + \frac{1}{4}\lambda(\pi^a \pi^a + \sigma^2)^2 \ . \tag{2.52}$$

We again see that only the σ field has a quadratic term so this gives rise to a massive particle, while quantising the π^a will give $N-1$ massless particles. These are the Goldstone bosons from spontaneous symmetry breaking.

Although the π^a fields are massless, they still appear in the potential (2.52), just in higher order terms. This is in contrast to the case with $U(1)$ symmetry where the potential didn't depend on the Goldstone field $\theta(x)$. There's no mystery here: it's because we've made no attempt to pick our fields to parameterise the vacuum moduli space $\mathcal{M}_0$. Instead, the $\pi^a(x)$ fields are just linear displacements away from the vacuum, and if you move away linearly from a point in $\mathcal{M}_0$, you eventually end up climbing the potential.

To do better, we could write our fields as something akin to the polar ansatz (2.43). Alternatively, if we're at low energies so that we care only about the dynamics of the Goldstone bosons, and not about their interactions with massive excitations, then we could restrict ourselves to $\mathcal{M}_0$ by insisting that (2.48) is obeyed everywhere, meaning

$$(\pi^a)^2(x) + (\phi^N)^2(x) = v^2 \ . \tag{2.53}$$

We could use this to eliminate $\phi_N(x)$ in our original action (2.47). By construction, the potential term vanishes completely and we're left just with kinetic terms for the Goldstone modes

$$S = \int d^4x \ \frac{1}{2}\left(\partial_\mu \vec{\pi} \cdot \partial^\mu \vec{\pi} + \frac{(\vec{\pi} \cdot \partial_\mu \vec{\pi})(\vec{\pi} \cdot \partial^\mu \vec{\pi})}{v^2 - \vec{\pi} \cdot \vec{\pi}}\right) \ . \tag{2.54}$$

We see that the Goldstone modes now have rather non-trivial interactions between themselves, but these interactions are entirely kinetic. To get a sense for what the action (2.54) is telling us, let's restrict to $N = 3$. In this case, the constraint (2.53) can be solved by the usual polar coordinates on $\mathbb{R}^3$,

$$\pi^1 = v \sin\theta\, \cos\varphi\ , \quad \pi^2 = v\sin\theta\, \sin\varphi\ , \quad \phi^3 = v\cos\theta\ . \tag{2.55}$$

It's important to stress that these are polar coordinates on field space, and both $\theta(x)$ and $\varphi(x)$ are fields that parameterise the vacuum manifold $\mathcal{M}_0 = \mathbf{S}^2$. With this choice of parameterisation, the action (2.54) becomes

$$S = \int d^4x\ \frac{v^2}{2}\Big(\partial_\mu\theta\,\partial^\mu\theta + \sin^2\theta\,\partial_\mu\varphi\,\partial^\mu\varphi\Big)\ . \tag{2.56}$$

We recognise the metric $ds^2 = d\theta^2 + \sin^2\theta d\varphi^2$ on $\mathbf{S}^2$ hiding within this action. More generally, any choice of parameterisation of the constraint (2.53) will give an action for the Goldstone bosons that takes the schematic form

$$S = \int d^4x\ \frac{1}{2} g_{ab}(\pi)\partial_\mu\pi^a\,\partial^\mu\pi^b \tag{2.57}$$

with g_{ij} the round metric on $\mathcal{M}_0$. Actions of this kind, where the fields are themselves coordinates on some manifold $\mathcal{M}$ are known as *non-linear sigma models*. In this context, the manifold $\mathcal{M}$ is sometimes called the *target space*, because the fields $\pi_a(x)$ are maps from spacetime (which is $\mathbb{R}^{1,3}$ for us) to the target manifold $\mathcal{M}$.

Non-linear sigma models like (2.57) are non-renormalisable. That means that they don't make sense up to arbitrarily high energy scales. But that's entirely reasonable! The sigma model (2.57) is constructed so that it describes only the very low energy physics. As we reach energies or order $E \sim \sqrt{\lambda}v$, we will start to be able to climb up the hills of the potential and out of the vacuum manifold $\mathcal{M}_0$. The original theory (2.47) provides a renormalisable, UV completion of the non-linear sigma model.

The origins of the name "sigma model" are somewhat farcical. It comes from the original paper of Gell-Mann and Lévy who did a calculation similar to the one above, eliminating the field $\sigma(x)$ (which, recall, is related to $\phi_N(x) = v + \sigma(x)$) and then naming the resulting Lagrangian after the field they got rid off! We'll see what Gell-Mann and Lévy did, and what the $\sigma(x)$ field describes in our world, when we come to discuss aspects of chiral symmetry breaking in QCD in section 3.

2.2.2 Goldstone's Theorem in Classical Field Theory

With these examples under our belt, we can now look at the general case. We will do this twice: once from the perspective of the classical theory, then again in the quantum theory.

We start classical. Consider a theory with a bunch of scalar fields, which we collectively denote as ϕ, transforming in some representation of a global symmetry group G. We will take G to a be Lie group, so we're dealing with continuous symmetries rather than discrete symmetries.

These fields experience a potential $V(\phi)$ which has some space of minima that define the vacuum manifold of the theory:

$$\mathcal{M}_0 = \Big\{ \phi_0 \mid V(\phi_0) = V_{\rm min} \Big\} \ . \tag{2.58}$$

If the ground state is unique – in which case we will assume that it sits at $\phi_0 = 0$ – then $\mathcal{M}_0$ is just a single point and we're back to the usual story in which the symmetry is realised only on excited states.

The more interesting situation is when ϕ_0 is not unique, In this case, acting with some elements of G will typically move us from one point in $\mathcal{M}_0$ to another. Indeed, the generic situation is that all points in $\mathcal{M}_0$ can be reached by a symmetry transformation, meaning that if we take two points $\phi_0, \phi_0' \in \mathcal{M}_0$, then there is a $g \in G$ such that

$$\phi_0' = g\phi_0 \ . \tag{2.59}$$

We can see this, for example, in the $O(N)$ model described above where $\mathcal{M}_0 = \mathbf{S}^{N-1}$ and you can always rotate from one point on the sphere to any other.

While some elements of G will move us around $\mathcal{M}_0$, other elements leave the point ϕ_0 unchanged. It's useful to define the concept of the *stability group* H. If we sit at some point $\phi_0 \in \mathcal{M}_0$, then the group H is defined to be those elements of G which don't change ϕ_0,

$$H = \Big\{ h \in G \mid h\phi_0 = \phi_0 \ \Big\} \ . \tag{2.60}$$

The stability group H defined above depends on the choice of $\phi_0 \in \mathcal{M}_0$. Happily, however, if we pick a different point $\phi_0' \in \mathcal{M}_0$ then we will find ourselves with a stability group H' that is isomorphic to H. This is simple to show: if $\phi_0' = g\phi_0$ then then for each $h \in H$ we can construct $h' = ghg^{-1} \in H'$.

Again, we can use the $G = O(N)$ model as an example. For any point in $\mathcal{M}_0 = \mathbf{S}^{N-1}$, the stability group is $H = O(N-1)$. The way in which $O(N-1)$ is embedded in $O(N)$ depends on where we sit in $\mathcal{M}_0$. For example, if we sit in the vacuum $\phi_i = (0, 0, \ldots, v)$ then the surviving $O(N-1)$ resides in the upper-left block of the $N \times N$ matrix, while if we sit in the vacuum $\phi_i = (v, 0, \ldots, 0)$ then $O(N-1)$ resides in the lower-right block. But, wherever we sit, there is always an $O(N-1)$ subgroup that survives.

We say that the group G is *spontaneously broken* to the group H. We usually write this as $G \to H$. The field ϕ is what, in statistical physics, we call an *order parameter* for the symmetry G: its value in the ground state – either zero or non-zero – provides a litmus test for whether the symmetry G is broken. The vacuum manifold $\mathcal{M}_0$ can then be identified as the coset space

$$\mathcal{M}_0 \cong G/H \ . \tag{2.61}$$

Here the coset G/H is defined to be the set of equivalence classes, with $g_1 \sim g_2$ if there exists an $h \in H$ such that $g_1 = hg_2$.

Now we're in a position to state the main result[3]:

Goldstone's Theorem: If a global, continuous symmetry G is spontaneously broken to H then the number of massless Goldstone bosons is given by

$$\dim(G/H) = \dim G - \dim H \ . \tag{2.62}$$

In light of the identification (2.61), you can think of these Goldstone bosons as the modes that fluctuate along the vacuum manifold $\mathcal{M}_0$.

Returning, briefly, to our $O(N)$ model, the sphere can be viewed as the coset $\mathbf{S}^{N-1} = O(N)/O(N-1)$. We can do some simple counting. We have $\dim O(N) = \frac{1}{2}N(N-1)$ so $\dim O(N) - \dim O(N-1) = N-1 = \dim \mathbf{S}^{N-1}$.

Proof: The proof of Goldstone's statement is really just a matter of turning our intuition into some equations. Suppose that ϕ sits in a representation R of the symmetry group G. We'll denote the components of ϕ as ϕ^a with $a = 1, \ldots, \dim R$.

Consider how ϕ shifts under an infinitesimal symmetry transformation, $g\phi = \phi + \delta\phi$. If we denote the generators of G in the representation R as $(T^A)^a{}_b$, with $A = 1, \ldots, \dim G$, then we have

$$\delta\phi^a = i\alpha^A (T^A)^a{}_b \phi^b \tag{2.63}$$

with α^A infinitesimal parameters. We know that G is a symmetry of our theory which means, among other things, that the potential must satisfy $V(g\phi) = V(\phi)$. So, for an infinitesimal transformation,

$$V(\phi + \delta\phi) - V(\phi) = i\alpha^A \frac{\partial V}{\partial \phi^a} (T^A)^a{}_b \phi^b = 0 \ . \tag{2.64}$$

[3] Both the classical and quantum versions of Goldstone's theorem were first proved by Goldstone, Salam and Weinberg in a classic 1962 paper entitled "Broken Symmetries". The proof was prompted by specific examples that had been explored by Nambu and by Goldstone.

We differentiate with respect to ϕ^b to find

$$\left[\frac{\partial V}{\partial \phi^a}(T^A)^a{}_b + \frac{\partial^2 V}{\partial \phi^a \partial \phi^b}(T^A)^a{}_b \phi^b\right] = 0 \tag{2.65}$$

where we've stripped off the α^A on the grounds that they are arbitrary parameters and so this expression must hold for each $A = 1, \ldots, \dim G$. Now we evaluate the result on a ground state ϕ_0. The first term disappears because ϕ_0 is a minimum of the potential and we're left with

$$\left.\frac{\partial^2 V}{\partial \phi^a \partial \phi^b}\right|_{\phi_0} (T^A \phi_0)^b = 0 \quad \text{for} \quad A = 1, \ldots, \dim G\ . \tag{2.66}$$

We recognise the second derivative of the potential as the mass matrix $M^2_{ab} = \partial V/\partial \phi^a \partial \phi^b$; the eigenvalues of this matrix are the physical masses. The result (2.66) is telling us that the mass matrix potentially has a bunch of zero eigenvalues, one for each eigenvector $(T^A\phi_0)^b$.

The "potentially" in the sentence above is there because it may be that the would-be eigenvector $(T^A\phi_0)^b$ actually vanishes. Indeed, this is clearly the case if $\phi_0 = 0$. That's as it should be: if $\phi_0 = 0$ then the symmetry is unbroken and there's no reason to generically expect massless modes. However, even when $\phi_0 \neq 0$, there will be some generators – let us call them $\tilde{T}^A$ – that annihilate the ground state,

$$\tilde{T}^A \phi_0 = 0\ . \tag{2.67}$$

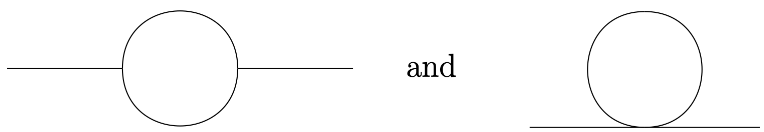
These are precisely the generators of the unbroken stability group H and so there are $\dim H$ of them. We will denote the generators orthogonal to $\tilde{T}^A$ as R^α, with $\alpha = 1, \ldots, \dim(G/H)$. Here, orthogonality means that they obey $\mathrm{Tr}\,(\tilde{T}^A R^\alpha) = 0$. Each of these generators gives a unique eigenstate $(R^\alpha \phi)_b$, and hence a massless mode. We see that there are at least $\dim(G/H)$ massless particles. These are the Goldstone bosons. $\square$

2.2.3 Goldstone's Theorem in Quantum Field Theory

The quantum version of Goldstone's theorem has much more teeth than its classical counterpart. This is not because the theorem itself is very much different – as we'll see, it really involves all the same ingredients that we've seen above, just adapted to life in a Hilbert space. Instead, the importance of the result is due to the environment in which the theorem operates.

In classical field theory, there's no difficulty in writing down a theory for a massless scalar. You literally just need to set $m^2 = 0$ in the potential. So while it's certainly interesting that spontaneous symmetry breaking gives us a mechanism for generating massless scalars, they're not such rare beasts.

But the story is very different for interacting quantum field theories. There, massless scalars (and, indeed scalars that are just "light" in some sense) are very hard to come by. This is because the physical mass is not just the m^2 that you write down in the Lagrangian. Instead, the mass of a scalar picks up extra contributions from the cloud of other fields that accompany the particle. These are captured, at one loop, by Feynman diagrams like this:



Here the external legs are the scalars, while the particle running in the loop is anything that the scalar interacts with, including itself. These diagrams contribute to the mass renormalisation of the scalar and, crucially, are quadratically divergent. Physically, it means that quantum corrections push the mass of a scalar particle up to the UV-cut off of the theory, Λ_{UV}.

The upshot of this is that, if you write down a Lagrangian with $m^2 = 0$, then it won't describe a quantum scalar particle with physical mass zero. Instead, after renormalisation, it will describe a scalar with physical mass $m^2 \sim \Lambda_{UV}^2$. (In some cases, Λ_{UV} may be some higher energy scale in the theory, rather than the UV-cut off. For example, in QCD we'll see that the masses of scalar mesons typically sit at a scale known as Λ_{QCD}.) If you want to write down, say a ϕ^4 theory that describes a massless scalar then you will need to tune the mass in the Lagrangian (the so-called "bare mass") to be $m^2 \sim -\Lambda_{UV}^2$, with a coefficient that precisely cancels the contributions from quantum corrections. This is known as *fine tuning* and it is generally agreed to be as tasteless as it sounds. (This same idea also arises in statistical physics, where the mass term is associated to the deviation from a critical temperature. In this case, the fine tuning is physical because you get to turn the temperature up and down at will.)

None of this means that there is some flaw in quantum field theory: instead it's capturing the right physics. Quantum field theories tend not to have massless, or indeed, light, scalar fields. Their mass is typically pushed up to some cut-off scale. This is not true of fermions, which suffer only a logarithmic correction to their mass.

This can be traced to the fact that fermions have an extra chiral symmetry when they are massless that protects their mass from being renormalised.

All of this means that things are interesting when you come across a physical system that does have a massless, or inordinately light, scalar field. If you find such a light scalar then there should be a reason why the arguments presented above fail. In most (but, famously, not all!) cases, that reason is Goldstone's theorem. Spontaneous symmetry breaking provides a robust mechanism to naturally deliver genuinely massless scalars, whose mass is protected against any corrections from renormalisation. And, as we mentioned at the beginning of this section, it is a mechanism that is employed over and over again by nature, from magnets, to phonons to, as we shall see later, pions.

Before we turn to prove Goldstone's theorem in the context of quantum field theory, it's worth commenting on the "famously, not all" remark above. This is a nod to the Higgs boson. It is not particularly light, weighing in at $m_H \approx 126$ GeV. But if we believe that quantum field theory continues to hold at scales significantly higher than m_H, we should ask why the mass of the Higgs boson hasn't been pushed up to higher scales. Or, in other words, why don't the simple arguments that we sketched above apply to the Higgs boson? We don't know the answer to this question. This is known as the *hierarchy problem*.

Broken Symmetries Acting on Hilbert Space

With this preamble in place, we can now see how Goldstone's theorem manifests itself in quantum field theory. We won't work with Lagrangians, or restrict ourselves to perturbation theory. Instead, all the physics can be seen in how symmetries act on the Fock space of particles.

By Noether's theorem, any continuous symmetry G has an associated set of currents J^A_μ, with $A = 1, \ldots, \dim G$. From these we can construct the conserved charges

$$Q^A = \int d^3x \; J^A_0 \; . \tag{2.68}$$

One of the lovely features of quantum mechanics (or, indeed, the Hamiltonian version of classical mechanics) is that these charges enact what we might call the "inverse Noether theorem". This means that, given a conserved charge, you can always reconstruct the associated symmetry. This follows from the fact that the charge is the generator of the symmetry, with any operator $\mathcal{O}$ undergoing the infinitesimal transformation

$$\delta_A \mathcal{O} = i[Q^A, \mathcal{O}] \; . \tag{2.69}$$

Comparing to our classical result (2.63), we see that our scalar fields ϕ^a transform as

$$[Q^A, \phi^a] = (T^A)^a{}_b \phi^b \ . \tag{2.70}$$

These are exact operator relations in the quantum theory.

In the classical theory, we saw that ϕ is an order parameter for the symmetry G. The same is true in the quantum theory, although strictly we should talk about the *vacuum expectation value* (or *vev*) of ϕ, as the order parameter,

$$\langle \phi \rangle = \langle \Omega | \phi | \Omega \rangle \tag{2.71}$$

where $|\Omega\rangle$ is the vacuum of the full, interacting theory. If $\langle \phi \rangle \neq 0$ then we say that ϕ *condenses*, a term taken from statistical physics. From (2.70), we have

$$\langle \Omega | [Q^A, \phi^a] | \Omega \rangle = (T^A)^a{}_b \langle \phi^b \rangle \neq 0 \ . \tag{2.72}$$

But this can only be true if

$$Q^A |\Omega\rangle \neq 0 \quad \text{for some } A \ . \tag{2.73}$$

This is what it means for a symmetry to be spontaneously broken in quantum field theory: the symmetry generators do not annihilate the vacuum.

Actually, there's a small caveat that I need to mention here. If we have $Q^A|\Omega\rangle = |\Omega\rangle$ then the commutator does vanish: $\langle \Omega | [Q^A, \phi^a] | \Omega \rangle = 0$. This kind of action on the ground state means that the symmetry is *unbroken* because, when exponentiated, we have $e^{i\alpha Q^A}|\Omega\rangle = e^{i\alpha}|\Omega\rangle$, but just changing the phase of a state in quantum mechanics is that the same as leaving the state invariant. So the statement $Q^A|\Omega\rangle \neq 0$ in (2.73) should be better written as $Q^A|\Omega\rangle \neq c|\Omega\rangle$ for some $c \in \mathbb{C}$.

For any symmetry generator, broken or unbroken, we have $[Q^A, H] = 0$ so (2.73) is really telling us that, whenever the symmetry is broken, the vacuum is degenerate. Said slightly differently, in quantum field theory every different choice of $\langle \phi \rangle$ corresponds to a different vacuum of the theory.

Conversely, if $\langle \phi \rangle = 0$ then, from (2.73), we see that the vacuum is annihilated by the symmetry generators: $Q^A|\Omega\rangle = 0$. This is the more familiar case in which the symmetry is unbroken. Excitations above the vacuum then sit in multiplets of G.

When a symmetry is spontaneously broken, the excitations above the vacuum no longer sit in multiplets of the full symmetry group G. To see this, suppose that we have two fields, ϕ^1 and ϕ^2, that are related by a symmetry so there is some conserved charge such that $[Q,\phi^1] = \phi^2$. We can consider excitations of the vacuum by the creation operators associated to ϕ^1, heuristically $|1\rangle = a_1^\dagger|\Omega\rangle$, and similar excitations associated to ϕ^2, $|2\rangle = a_2^\dagger|\Omega\rangle$. We then have

$$|2\rangle = a_2^\dagger|\Omega\rangle = [Q,a_1^\dagger]|\Omega\rangle = Q|1\rangle - a_1^\dagger Q|\Omega\rangle \; . \tag{2.74}$$

We see that the symmetry generator does relate $|1\rangle$ and $|2\rangle$ but only if $Q|\Omega\rangle = 0$. When the symmetry is spontaneously broken, so $Q|\Omega\rangle \neq 0$, the two states $|1\rangle$ and $|2\rangle$ can have different properties. For example, they may have different energies.

So far, we haven't described where the Goldstone bosons come from. Following our classical intuition, we expect them to correspond to fluctuations along the directions of broken symmetry. And that's indeed the case. For each broken symmetry generator, we construct states

$$|\pi^A(\mathbf{p})\rangle \sim \int d^3x \; e^{i\mathbf{p}\cdot\mathbf{x}} J_0^A(x)|\Omega\rangle \; . \tag{2.75}$$

These states carry 3-momentum $\mathbf{p}$. Moreover, in the limit of vanishing momentum, we have

$$\lim_{\mathbf{p}\to 0} |\pi^A(\mathbf{p})\rangle \sim Q^A|\Omega\rangle. \tag{2.76}$$

For those generators that are spontaneously broken, the state $Q^A|\Omega\rangle \neq 0$ has the same energy as the original vacuum $|\Omega\rangle$ because $[Q^A,H] = 0$. This is the statement that the Goldstone boson $|\pi^A(\mathbf{p})\rangle$ has energy $E \to 0$ as $\mathbf{p} \to 0$. In other words, the Goldstone boson is massless.

None of the arguments above rely on perturbation theory: they are all exact statements about the interacting quantum field theory. This means that if we were to write down Lagrangians for these Goldstone bosons then they must remain massless, even after taking into account one-loop effects and so on. In operational terms, this happens because the Goldstone bosons have only derivative couplings.

The argument above is not completely rigorous, not least because $Q|\Omega\rangle$ suffers from divergences and doesn't strictly exist in the Fock space. A better, but more formal, argument uses the Källén-Lehmann spectral decomposition. You can read about this in Volume II of Weinberg's book.

The View From the Effective Potential

There is an alternative proof of Goldstone's theorem in quantum field theory that follows much more closely the classical proof that we saw previously. We first need to review some basic facts about generating functions in quantum field theory. The generating function for connected correlation functions is

$$e^{iW[J]} = \int \mathcal{D}\phi \; e^{i\int d^4x \; (\mathcal{L}(\phi)+J\phi)} \; . \tag{2.77}$$

Here $J(x)$ is a source for ϕ and differentiating $W[J]$ successively with respect to $J(x)$ gives the connected correlation functions. In particular, the expectation value of $\phi(x)$ is given by

$$\frac{\delta W[J]}{\delta J(x)} = \langle \Omega|\phi(x)|\Omega\rangle = \phi_{\rm cl}(x) \; . \tag{2.78}$$

In the absence of a source, Lorentz invariance implies that $\phi_{\rm cl}$ is just a number, and coincides with the vev (2.71) that we introduced previously. But, if we turn on a spatially varying source $J(x)$, then the function $\phi_{\rm cl}(x)$ will respond accordingly.

The Legendre transform of $W[J]$ is known as the *one-particle irreducible* (or *1PI* for short) effective action,

$$\Gamma[\phi_{\rm cl}] = W[J] - \int d^4x \; J(x)\phi_{\rm cl}(x) \; . \tag{2.79}$$

As in other examples of Legendre transforms, we should use (2.78) to replace $J(x)$ with $\phi_{\rm cl}(x)$ in the 1PI effective action. We can always return to $W[J]$ (assuming certain convexity properties) using

$$\frac{\delta\Gamma[\phi_{\rm cl}]}{\delta\phi_{\rm cl}(x)} = -J(x) \; . \tag{2.80}$$

The 1PI effective action is not, in general, the same thing as the more physical Wilsonian effective action that we get by integrating out high energy modes to find a description of the low energy physics. Taking derivatives of $\Gamma[\phi_{\rm cl}]$ generates the 1PI Green's functions. In particular, the two derivative term gives the inverse propagator

$$\frac{\delta^2\Gamma}{\delta\phi_{\rm cl}(x)\delta\phi_{\rm cl}(y)} = \Delta^{-1}(x-y) \; . \tag{2.81}$$

In general, $\Gamma[\phi_{\rm cl}]$ can be expressed in terms of a derivative expansion,

$$\Gamma[\phi_{\rm cl}] = \int d^4x \; \Big(- V_{\rm eff}(\phi_{\rm cl}) + \frac{1}{2}Z(\phi_{\rm cl})\partial_\mu\phi_{\rm cl}\partial^\mu\phi_{\rm cl} + \ldots \Big) \; . \tag{2.82}$$

For our purposes, we're interested only in spatially homogeneous configurations, so we can ignore the derivative terms and the 1PI effective potential becomes

$$\Gamma[\phi_{\rm cl}] = -\mathcal{V} V_{\rm eff}(\phi_{\rm cl}) \tag{2.83}$$

where $\mathcal{V}$ is the (admittedly infinite, but actually irrelevant) volume of spacetime. Restricted to constant configurations, the second derivative of $\Gamma[\phi_{\rm cl}]$ is just the mass matrix, but now for the physical masses as opposed to the classical, bare masses

$$\frac{\partial^2 V_{\rm eff}}{\partial\phi_{\rm cl}\partial\phi_{\rm cl}} = \Delta^{-1}(0)\ . \tag{2.84}$$

Spontaneous symmetry breaking occurs when we have $\phi_{\rm cl} \neq 0$ even when $J = 0$. From (2.80), this translates into the familiar requirement that

$$\phi_{\rm cl} \neq 0 \quad \text{at} \quad \frac{\partial V_{\rm eff}}{\partial \phi_{\rm cl}} = 0\ . \tag{2.85}$$

Now we may rerun all the arguments of section 2.2.2, but for the effective potential $V_{\rm eff}(\phi)$ rather than the classical potential $V(\phi)$ to again arrive at (2.66),

$$\frac{\partial^2 V_{\rm eff}}{\partial\phi^a_{\rm cl}\,\partial\phi^b_{\rm cl}}(T^A\phi_0)^b = 0\ . \tag{2.86}$$

As in the classical argument, this is telling us that the mass matrix has a number of zero eigenvalues (Equivalently, the propagator Δ has poles at $\mathbf{p} \to 0$.) There is one zero eigenvalue for each broken generator.

2.2.4 The Coleman-Mermin-Wagner Theorem

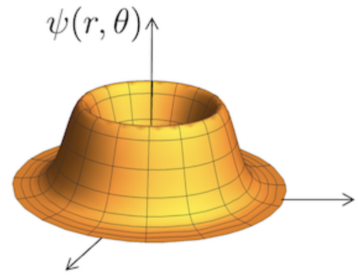
In all our discussions above, we assumed that spontaneous symmetry breaking actually takes place in the quantum theory. For example, we showed that *if* $\langle\phi\rangle \neq 0$ then the ground state must necessarily shift under a symmetry

$$Q|\Omega\rangle \neq 0\ . \tag{2.87}$$

But how do we know that this actually happens? In particular, there is some tension with what we know from our first courses on quantum mechanics.

Let's return to the simplest example of a Mexican hat potential (2.38), but now think of quantum mechanics, rather than quantum field theory. That means that we have a quantum particle moving in the potential.

It's challenging to write down the exact ground state wavefunction $\psi(r,\theta)$, but it's not difficult to get some idea of what it looks like: it will be peaked in the trough at $r = v$, and be fully delocalised in the angular θ direction. In other words, it will look something like the wavefunction shown in the figure. But, crucially, because the wavefunction spreads around the circle parameterised by θ, there is no spontaneous symmetry breaking.



This begs the question: why is quantum field theory different from quantum mechanics? Why do we expect spontaneous symmetry breaking in the former case, but not in the latter? A similar question arose when we discussed discrete symmetries and there we understood that quantum tunnelling through the barrier was suppressed by the infinite spatial volume. But here there's no barrier to tunnel through. Instead we have a manifold of ground states $\mathcal{M}_0$ and it feels like it should be easier for a wavefunction to spread over $\mathcal{M}_0$ than to tunnel through a barrier. In other words, it should be more difficult to spontaneously break continuous symmetries than to spontaneously break discrete symmetries.

And indeed it is. But in an interesting way. The key physics is captured by the following theorem:

Theorem: A continuous symmetry cannot be broken in quantum theories in $d = 0+1$ (i.e quantum mechanics) or $d = 1+1$ dimensions.

This theorem was first proven by Mermin and Wagner for certain spin chains, inspired by previous work by Hohenberg. The proof in the context of quantum field theory is due to Coleman[4]. We see that the story is different for discrete and continuous symmetries. A discrete symmetry can be spontaneously broken in spacetime dimensions $d = 1+1$ and higher, but for a continuous symmetry to be spontaneously broken we must be in $d = 2+1$ or higher.

[4]The original paper is from 1966, "Absence of Ferromagnetism or Anti-Ferromagnetism in One or Two-Dimensional Heisenberg Models" by Mermin and Wagner and, because of quirk of publication, appeared before the Hohenberg paper which motivated them: "Existence of Long-Range Order in One and Two Dimensions". Sidney Coleman's contribution is from 1973, in the concisely titled "There are no Goldstone Bosons in Two Dimensions".

Here we offer just a sketch of this theorem. In fact, the basic idea can already be seen in classical field theory. Things are simplest if we work in d-dimensional Euclidean space. Suppose that we have a massless scalar field ϕ with no potential. This means that we have a choice of what we call the vacuum and, for our purposes, we'll decide that $\phi = 0$ is the ground state. Now we excite this scalar field by introducing a delta function source at the origin. That means that we have to solve

$$\nabla^2 \phi = \delta(\mathbf{x}) \ . \tag{2.88}$$

This, of course, is the equation for the Green's function of the d-dimensional Laplacian. The solutions take the schematic form (ignoring overall coefficients)

$$\phi(\mathbf{x}) \sim \begin{cases} |x| & \text{for } d = 1 \\ \log|\mathbf{x}| & \text{for } d = 2 \\ 1/|\mathbf{x}|^{d-2} & \text{for } d \geq 3 \end{cases} \tag{2.89}$$

We see that for low dimensions, $d = 1$ and $d = 2$, exciting the scalar field at the origin means that it can no longer take the value $\phi = 0$ asymptotically. Any disturbance at the origin is still felt at $|\mathbf{x}| \to \infty$ where the field continues to grow. In contrast, in $d = 3$ and higher, the field is excited near the origin but then settles back down to $\phi \to 0$ as $|\mathbf{x}| \to \infty$.

The story above is classical. What happens in the quantum theory? We'll stick with the free massless scalar, and continue to work in Euclidean spacetime. Consider the two-point function $\langle \phi(\mathbf{x})\phi(\mathbf{y})\rangle$. We know from the lectures on Quantum Field Theory that this is given by the same Green's function as above, so

$$\langle \phi(\mathbf{x})\phi(\mathbf{y})\rangle \sim \begin{cases} |x - y| & \text{for } d = 1 \\ \log|\mathbf{x} - \mathbf{y}| & \text{for } d = 2 \\ 1/|\mathbf{x} - \mathbf{y}|^{d-2} & \text{for } d \geq 3 \end{cases} \tag{2.90}$$

Again, we see the infra-red divergence for $d = 1$ and $d = 2$. Roughly speaking, this is telling us that the wavefunction spreads over all values of ϕ in $d = 2$ dimensions, just as it does in $d = 1$ quantum mechanics. In both cases, there is no normalisable ground state.

A better way of saying this is that $\phi(\mathbf{x})$ is not a well defined operator in $d = 2$ dimensions. In particular, the correlation function $\langle \phi(\mathbf{x})\phi(\mathbf{y})\rangle \sim \log|\mathbf{x} - \mathbf{y}|$ is not positive for all $\mathbf{x} - \mathbf{y}$, which is one of the requirements of a QFT. However, although $\phi(\mathbf{x})$ is not a well-defined operator, its derivatives $\partial_\mu \phi(\mathbf{x})$ are. You can learn more about this 2d theory (which really only makes sense when ϕ is taken to be a periodic variable) in the lectures on String Theory.

No such problems arise for a massless scalar in $d \geq 3$ spacetime dimensions. Here, each value of $\langle\phi\rangle$ specifies a different ground state of the theory. Indeed, for this simple free theory, the massless ϕ field can be viewed as a Goldstone boson for the shift symmetry $\phi \to \phi + \text{constant}$.

As for the discrete symmetries discussed in Section 2.1, the existence of spontaneous symmetry breaking is due to the infinite volume of space. If we were to take our quantum field theory on a compact spatial manifold, then the long-time behaviour is the same as in quantum mechanics, and the wavefunction will again spread over field space, obviating spontaneous symmetry breaking.

2.3 The Higgs Mechanism

Goldstone's theorem tells us that when a continuous symmetry is spontaneously broken, it results in a massless boson. Here we would like to ask: what happens if that symmetry is gauged?

First, the very concept of a "spontaneously broken gauge symmetry" is a little misleading. As we've stressed, a gauge symmetry is merely a redundancy in the description of a system and there's no way that this redundancy can be "broken" or "lost". This linguistic issue notwithstanding, the physics underlying the spontaneous breaking of gauge symmetries is clear cut. First, there is no massless Goldstone boson. Second, the gauge boson gets a mass. We'll now see, in some detail, how this comes about.

2.3.1 The Abelian Higgs Model

We return to a complex scalar ϕ with the Mexican hat potential of Section 2.2. This time, however, we couple the scalar to a $U(1)$ gauge field. The action is

$$S = \int d^4x \, \Big(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mathcal{D}_\mu\phi^\dagger \mathcal{D}^\mu\phi - \frac{\lambda}{2}(|\phi|^2 - v^2)^2 \Big) \tag{2.91}$$

This is known as the *Abelian Higgs model*. The covariant derivative is $\mathcal{D}_\mu\phi = \partial_\mu\phi - ieA_\mu\phi$. Clearly the ground state sits at

$$|\phi|^2 = v^2 \ . \tag{2.92}$$

Previously, this meant that we had a vacuum manifold, $\mathcal{M}_0 = \mathbf{S}^1$, parameterised by the phase of ϕ. But now the $U(1)$ that takes us around the $\mathbf{S}^1$ is a gauge symmetry,

$$\phi \to e^{ie\alpha(x)}\phi \quad \text{and} \quad A_\mu \to A_\mu + \partial_\mu\alpha \tag{2.93}$$

and we know that field configurations that are related by gauge symmetries should be considered physically equivalent. This suggests that the gauge theory only has a single ground state, rather than a manifold of ground states. This, it turns out, is the right interpretation.

To see the physics, let's place ourselves in the classical vacuum $\phi = v$ and look at fluctuations that we parameterise as

$$\phi(x) = e^{i\theta(x)}\big(v + \sigma(x)\big) \ . \tag{2.94}$$

We then have

$$\mathcal{D}_\mu \phi = e^{i\theta}\big(\partial_\mu \sigma + i(v+\sigma)(\partial_\mu\theta - eA_\mu)\big) \ . \tag{2.95}$$

Substituting this into the action, and expanding out, we have

$$S = \int d^4x \ \left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \partial_\mu\sigma\partial^\mu\sigma + (v+\sigma)^2(\partial_\mu\theta - eA_\mu)(\partial^\mu\theta - eA^\mu) - V(\sigma)\right) \ .$$

with

$$V(\sigma) = \frac{\lambda}{2}\sigma^2(\sigma + 2v)^2 \ . \tag{2.96}$$

From this, we can read off the mass spectrum of the theory. First, the scalar σ is reasonably standard: it has a quadratic term that tells us its mass is

$$m_\sigma^2 = 2\lambda v^2 \ . \tag{2.97}$$

This is the same mass that we calculated for the global symmetry. Later, when we discuss electroweak theory, we will learn that an analogous particle is the Higgs boson.

More interesting is the other scalar field $\theta(x)$. In the absence of the gauge field, this was the Goldstone boson. But now that we've introduced the gauge field, we see something interesting: this field only appears in kinetic terms in the combination

$$\partial_\mu\theta - eA_\mu \ . \tag{2.98}$$

This allows us to eliminate the field $\theta(x)$ complete. We simply define a new gauge field, related to the first by the change of variables

$$A'_\mu = A_\mu - \frac{1}{e}\partial_\mu\theta \ . \tag{2.99}$$

This has the same field strength as A_μ, with $F_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu$. However, in contrast to A_μ, the new field A'_μ does not change under a gauge transformation since the usual shift $A_\mu \to A_\mu + \partial_\mu\alpha$ is now compensated by $\theta \to \theta + e\alpha$. Said slightly differently, you could also think of the change of variables to A'_μ as analogous to working in $\theta = 0$ gauge, known, in this context, as *unitary gauge*. Either way, the upshot is the same: the field $\theta(x)$ no longer appears in the action

$$S = \int d^4x\ \left(-\frac{1}{4} F_{\mu\nu}F^{\mu\nu} + \partial_\mu\sigma\partial^\mu\sigma + e^2(v+\sigma)^2 A'_\mu A'^\mu - V(\sigma) \right) \ . \tag{2.100}$$

We see that we've generated a mass term $e^2v^2A'_\mu A'^\mu$ for the gauge field. This is exactly the kind of term that is usually forbidden by gauge invariance. But such a term arises naturally when we spontaneously break the gauge symmetry and the photon gets a mass

$$m_\gamma^2 = 2e^2v^2 \ . \tag{2.101}$$

This is the *Higgs mechanism*.

There's some interesting interplay of degrees of freedom going on here. Massive spin 1 particles have three degrees of freedom. (This is just the $(2l+1)$-dimensional representation of the little group for $l = 1$.) But massless spin 1 particles have only two degrees of freedom, the two polarisation states. But it's clear where the extra degree of freedom came from because the photon absorbed the would-be Goldstone mode $\theta(x)$. This Goldstone boson breathes life into the longitudinal mode of the photon which is ordinarily killed by the constraints of gauge invariance.

Note that the mass of the Higgs boson (2.97) and the mass of the photon (2.101) have different parameteric dependence on the coupling constants. This means, among other things, that we could always just decouple the Higgs boson by taking $m_\sigma \to \infty$, leaving behind the massive photon at a finite mass m_γ. Given this, you might wonder why we needed all this palava with the Higgs boson. And, in fact, we really don't. We could always just couple the photon directly to the Goldstone mode θ, ignoring the radial mode σ. Said differently, we could just couple the photon to the sigma model with target space $\mathcal{M}_0 = \mathbf{S}^1$ which gives a massive photon and no Higgs boson. However, this option is less viable when we discuss the Higgs mechanism in non-Abelian theories because the corresponding sigma model is non-renormalisable and so should be viewed as an effective low energy theory, breaking down in the UV.

2.3.2 Superconductivity

We will later see that the Higgs mechanism plays a key role in the Standard Model. But there is a glorious unity to physics, and if nature finds a good trick to use in one context, she often recycles it elsewhere. So it is with the Higgs mechanism, which also provides a description of how superconductors work. In that context, it is referred to as the *Anderson-Higgs mechanism*[5].

Superconductivity is a phenomenon exhibited by many metals when they are cooled to a few degrees Kelvin. The metal undergoes a phase transition, and the electrical resistivity promptly plummets. At the same time, any magnetic fields are expelled.

The microscopic explanation for superconductivity is beyond the scope of these lectures. For what it's worth, an attractive coupling mediated by the phonon causes electrons to form an object known as a Cooper pair. For our purposes, all we need to know is that the resulting bound state is described by a complex scalar field ϕ that has charge $-2e$, with the -2 because it's formed of two constituent electrons.

In condensed matter physics, we more commonly work with the free energy, which describes the equilibrium properties of a system at finite temperature, rather than the Lagrangian which describes the zero temperature dynamics. But to avoid taking too much of a detour, here we give a Lagrangian description of superconductivity. This is almost identical to the Abelian Higgs model of the previous section, with just one small difference: the dynamics of the scalar field ϕ is non-relativistic. This means that we should work with the action

$$S = \int dt\, d^3x \, \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\phi^\dagger \mathcal{D}_t \phi - |\mathcal{D}_i \phi|^2 - \frac{\lambda}{2}(|\phi|^2 - v^2)^2 \right) \,. \tag{2.102}$$

In addition, there's an extra factor of -2 buried in the covariant derivatives: $\mathcal{D}_\mu \phi = \partial_\mu \phi + 2ieA_\mu \phi$. (On dimensional grounds, there should be a coefficient with dimension $(\text{mass})^{-1}$ in front of the gradient terms but I've set it to unity to ease comparison with the relativistic Abelian Higgs model (2.91).)

A non-relativistic complex scalar has just a single degree of freedom. (This is true because the kinetic term contains a first order time derivative and so $\phi^\dagger$ is the momentum conjugate to ϕ, rather than a separate degree of freedom.) This means that if we

[5]The history of the Higgs phenomenon is famously murky. Anderson's 1963 paper on superconductivity argues that the would-be Goldstone mode is no longer there and that the photon is gapped. These ideas were extended to the relativistic theory by Brout and Englert and, independently, by Peter Higgs. Only Higgs' paper mentions the existence of an additional massive particle, now called the *Higgs boson*, albeit in what appears to be an afterthought in the final paragraph of the paper. You can decide for yourself whether this was because the existence of the Higgs boson was obvious (as some of the authors later claimed) or because they didn't think to ask the question. Still, the mechanism for giving a photon mass should probably rightly be called the Anderson-Brout-Englert-Higgs mechanism. In line with much of the particle physics community, we chose to unfairly shorten this to simply "Higgs". Meanwhile the term Higgs boson, for the particle, seems more appropriate.

quantise (2.102), we will find a massive photon, but the would-be Higgs boson (what we called σ in the relativistic theory) is missing.

We can read off the charge density and current from the coupling $A_\mu J^\mu$. The charge density is

$$J^0 = -2e|\phi|^2 \ . \tag{2.103}$$

In the ground state, we have the condensation $|\phi|^2 = v^2$, so the Cooper pairs form a constant background electric charge. (In a real system, this is compensated by the positive electric charge of the underlying lattice of ions.) Meanwhile, assuming that $|\phi|^2 = v^2$, the electric current is

$$\mathbf{J} = 4ev^2 \left(\nabla\theta - 2e\mathbf{A}\right) \ . \tag{2.104}$$

Here, as in the previous section, $\theta(x)$ is the phase of $\phi(x)$. The expression (2.104) is known as the *supercurrent*. It is sometimes denoted as $\mathbf{J}_s$ to distinguish it from the normal current carried by electrons.

Resistance is Futile

The signature of a superconductor is that it conducts electricity without resistance. This follows immediately from the equation of motion for $\phi^\dagger$,

$$i\mathcal{D}_0\phi = -\mathcal{D}^2\phi + \frac{\partial V}{\partial \phi^\dagger} \ . \tag{2.105}$$

In the lowest energy state, the charge density $|\phi|^2$ is constant. But the phase can vary. Indeed, from (2.104), we see that a spatially varying phase $\nabla\theta \neq 0$ means that an electric current flows.

Suppose that we look at such a configuration with $|\phi|^2 = v^2$. Then the complex equation of motion (2.105) splits into real and imaginary parts, which are

$$\dot{\theta} - 2eA_0 = \frac{1}{(4ev^2)^2}\,\mathbf{J}^2 \quad \text{and} \quad \nabla\cdot\mathbf{J} = 0 \ . \tag{2.106}$$

To see the relevant physics, it's simplest to restrict to the case where $\mathbf{J}$ is constant in space so that $\nabla\mathbf{J}^2 = 0$. Then, taking the time derivative of the (2.104), we have

$$\frac{d\mathbf{J}}{dt} = 4ev^2\left(\nabla\dot{\theta} - 2e\dot{\mathbf{A}}\right) = 2(2ev)^2| \left(-\nabla A_0 - \dot{\mathbf{A}}\right) = 2(2ev)^2\mathbf{E} \ . \tag{2.107}$$

This is the *first London equation*. It tells us that an electric field acts to accelerate the current, rather than to maintain the current. But that's not what usually happens

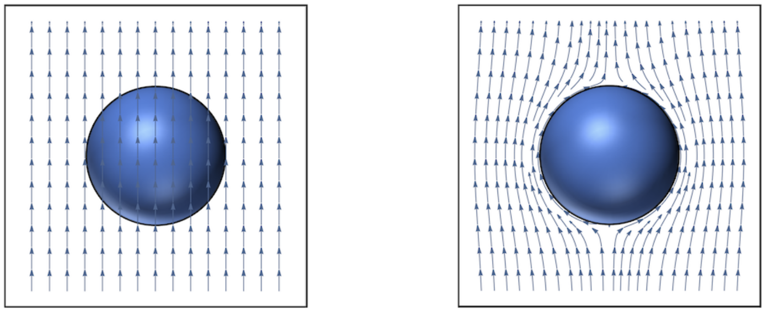


Figure 4. A constant magnetic field can pass through a normal metal, as shown on the left. But when the metal becomes superconducting, as shown on the right, the magnetic field is expelled, a phenomenon known as the Meissner effect.

in a conductor. Usually, a constant electric field induces a constant current. That's what the famous Ohm's law equation $V = IR$ says. But the resistance R in a normal conductor is due to friction terms, and the London equation (2.107) is telling us that a superconductor has vanishing resistance, $R = 0$.

Meissner Effect

Superconductors don't like magnetic fields very much. If you try to force a magnetic field through a superconductor, then it will resist. This is known as the *Meissner effect*, or sometimes as the *Meissner-Ochsenfeld effect*. A cartoon of this is shown in Figure 4. It has the dramatic consequence that a superconductor, placed above a magnet, is repelled and can levitate in mid-air.

At heart, the Meissner effect arises because the photon gets a mass. The term $\sim v^2\mathbf{A}\cdot\mathbf{A}$ in the action ensures that it is energetically costly to turn on a magnetic field.

We can see this more quantitatively from the form of the supercurrent (2.104). If we take the curl of both sides, we find

$$\nabla \times \mathbf{J} = -2(2ev)^2\mathbf{B} \ . \tag{2.108}$$

This is the *second London equation*. We can compare it to Ampére's law, $\nabla\times\mathbf{B} = \mu_0\mathbf{J}$. Taking the curl, and using $\nabla\times\nabla\times\mathbf{B} = -\nabla^2\mathbf{B}$ (because $\nabla\cdot\mathbf{B} = 0$), we find that the magnetic field inside a superconductor obeys the Helmholtz equation

$$\nabla^2\mathbf{B} = \frac{1}{\lambda^2}\mathbf{B} \quad \text{with} \quad \lambda^2 = \frac{1}{2(2ev)^2} \ . \tag{2.109}$$

Here λ is the *penetration depth*, a length scale equal to the inverse mass of the photon, $\lambda = 1/m_\gamma$. (The factor of 4 difference with (2.101) can be traced to the fact that, for superconductors, we're dealing with a field with charge $-2e$ rather than e.)

To see why the penetration depth gets it name, we can solve this equation for a constant magnetic field of the form

$$\mathbf{B} = (0, 0, B(z)) \ . \tag{2.110}$$

Suppose that the superconductor fills half of space, say the region with $z > 0$. We set up a constant magnetic field $\mathbf{B} = (0, 0, B_0)$ in the outside region $z < 0$ and ask what becomes of it when it enters the superconductor. There are two solutions to (2.109), but only the decaying one is physical. We find that the magnetic field drops off exponentially quickly inside the superconductor,

$$B(z) = B_0 e^{-z/\lambda} \ . \tag{2.111}$$

This is the Meissner effect: the superconductor does not suffer a magnetic field inside. In most superconductors, $\lambda \approx 10^{-8}$ to 10^{-9} m. This is what allows superconducting materials to levitate above magnets: the magnetic field can't penetrate the superconductor, and has to go around as shown in Figure 4. This squeezes the magnetic field lines which costs energy, making it energetically preferable for the superconductor to remain magically suspended in space, rather than falling like other materials that have more respect for gravity.

Vortices

There's no such thing as an immovable object. If you push hard enough, by cranking up the magnetic field, then the superconductor will eventually relent and let it pass. But the way it does this is interesting.

This follows because of a novel solution to the equations of motion of the action (2.102) known as a *vortex*. (This is also a solution to the relativistic Abelian Higgs model (2.91).) The vortex solution is time-independent, and extends along one spatial direction – say the z-direction – as a string-like object. To this end, we will look for solutions with $\partial_0 = \partial_3 = 0$ as well as $A_0 = A_3 = 0$.

It turns out that no closed form solution to the resulting equations of motion is known (although it is not hard to construct numerically). So rather than try to solve the equations directly, we will instead argue that such a solution must exist. The argument involves a little simple topology.

Consider the (x, y)-plane at $z = 0$. We will work with 2d polar coordinates $x + iy = re^{i\varphi}$. The trick is to look for solutions such that, for any curve C around the origin, we have

$$\oint_C \nabla\theta \cdot d\mathbf{x} \neq 0 \ . \tag{2.112}$$

Our first task is to understand what this means. Usually, the integral of a total derivative is zero, but in the present case there's an opportunity for something more interesting to happen. This is because the field θ started life as a phase of our scalar ϕ and, as such, is periodoc, taking values $\theta \in [0, 2\pi)$. For a periodic field θ, the line integral $\oint_C \nabla\theta \cdot d\mathbf{x}$ counts the number of times that θ winds as we traverse the curve C.

For example, if the curve C is parameterised by a coordinate $\varphi \in [0, 2\pi)$ then we could consider field configurations of the form $\theta = k\varphi$. Because θ must be single-valued, this only makes sense for $k \in \mathbb{Z}$ which is acceptable because $\theta = 0$ is equivalent to $\theta = 2\pi$. This, in turn, means that the integral (2.112) is necessarily quantised,

$$\oint_C \nabla\theta \cdot d\mathbf{x} = \int_0^{2\pi} d\varphi \ \frac{d\theta}{d\varphi} = 2\pi k \quad \text{with} \quad k \in \mathbb{Z} \ . \tag{2.113}$$

This quantisation doesn't happen because of anything to do with quantum mechanics. Instead, it's a quantisation imposed upon us by simple topological configurations.

Let's look for configurations in which the phase θ has winding (2.112). If this configuration is to have finite energy (per unit length) then, asymptotically, we must have $\mathcal{D}_i\phi \to 0$. This tells us that

$$\oint_C \nabla\theta \cdot d\mathbf{x} = 2e \oint_C \mathbf{A} \cdot d\mathbf{x} = 2e \int d^2x \ B_3 = 2e\Phi \tag{2.114}$$

with Φ the magnetic flux through the plane. We see that the quantisation of the winding translates into a quantisation of the allowed magnetic flux

$$\Phi = \frac{2\pi}{2e} k \quad \text{with} \quad k \in \mathbb{Z} \ . \tag{2.115}$$

I've not cancelled the factors of 2 here to stress the fact that, by measuring the minimal unit of flux, with $k = \pm 1$, you can determine that the current is carried by particles of charge $\pm 2e$, rather than the electron charge $-e$. (Indeed, this was one of the first experiments to confirm the charge of the condensate in a superconductor.)

The quantisation of winding means that the field configurations in this theory split into distinct topological sectors, labelled by $k \in \mathbb{Z}$. Because this integer is determined by the asymptotic boundary conditions, there's no way that a field configuration in one topological sector can move smoothly into a configuration in another. This means that if we can find novel solutions to the equations of motion by minimising the energy (per unit length) in any given sector.

Let's think about how this works for the minimum winding $k = 1$. Because the winding number is quantised, it can't change gradually as we vary the radius of the contour C in (2.113). It must give the same value $k = 1$ for all choices of C. That's all fine until we get to the origin, at which point the phase θ gets something of an identity crisis because it's supposed to point in all directions at once. The only way out is to realise that θ is the phase of the field ϕ, and so there must be a point in the (x, y)-plane where $\phi = 0$ so that the phase is ill-defined. This means that whenever we have winding, there is necessarily a small region of non-superconducting phase, with $\phi = 0$, somewhere inside the contour C. That will be the region where it is energetically preferable for the flux Φ in (2.115) to penetrate.

We can get an estimate for the size of the region over which the condensate varies. For simplicity, we set $A_0 = \mathbf{A} = 0$ and restrict to time-independent configurations $\phi(x, y)$. Then the equation of motion (2.105) reads

$$\nabla^2\phi = \lambda\phi(|\phi|^2 - v^2) \ . \tag{2.116}$$

This equation contains a natural length scale ξ, given by

$$\xi^2 = \frac{1}{\lambda v^2} \ . \tag{2.117}$$

This is known as the *coherence length*. It is roughly equal to the inverse mass of the scalar (2.97) in the relativistic theory: $\xi = \sqrt{2}/m$. (That factor of $\sqrt{2}$ is just annoying convention.) The coherence length sets the scale over which the condensate ϕ is roughly zero (or, more precisely, exponentially small) in the vortex solution. In most superconductors, the coherence length is within a couple of orders of magnitude of the penetration depth, λ, the analogous quantity for the magnetic field.

We could put more meat on this discussion by explicitly solving the equations of motion for the gauge field and scalar. By making a suitable, rotationally invariant ansatz, you can reduce these equations to two, coupled ordinary non-linear differential equations. There is no solution in closed form, but it is straightforward to solve them numerically. A schematic picture of the resulting condensate and magnetic flux, as a

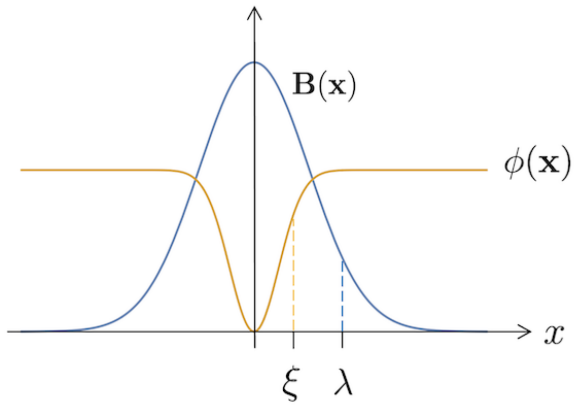


Figure 5. The spatial profile of the magnetic field and condensate for a vortex.

cut-through in the x-direction, is shown in Figure 5 in the case where $\lambda > \xi$, so the magnetic field spills out over the region where $\phi = 0$.

The discussion above took place in the $z = 0$ plane. But we can repeat the story as we move the contour C in the z-direction. The winding can't change, and so the region with $\phi = 0$ and magnetic flux necessarily extends in the z-direction. In other words, we have a magnetic flux tube. This is the *vortex*.

The fact that non-linear equations of motion have novel localised solutions like the vortex is interesting. In particular, the existence of this solution can be traced to the topological nature of the winding. The general name given to solutions of this kind is *soliton*.

For the story above, we restricted attention to the minimal $k = 1$ sector. What happens for higher $k \geq 2$ is also interesting and depends on the ratio of the two length scales ξ/λ. There are three possibilities:

- For $\xi > \sqrt{2}\lambda$, the scalar field ϕ spreads out further than the magnetic flux. But there is a general story that magnetic flux repels, while scalar fields attract. (For example, the Yukawa force is always attractive.) This means that two vortices will feel an attractive force, albeit one that is exponentially suppressed on scales $r \gg \xi$. This is what happens in a *Type I superconductor*.

  What actually happens in practice is that, if you apply a magnetic field to a Type I superconductor, then the whole material will transition to the normal, metallic phase at some critical magnetic field B_c. This means that you don't see vortices in this case.

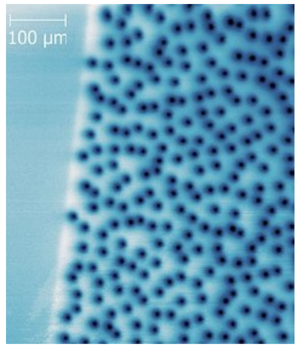


Figure 6. The Abrikosov vortex lattice, observed in the high temperature superconductor YBCO.

- For $\xi < \sqrt{2}\lambda$, the magnetic field spreads out further than the scalar field, as shown in Figure 5. In this case, two nearby vortices experience a repulsive force. This is known as a *Type II superconductor*.

  If you apply a magnetic field to a Type II superconductor then, initially, the superconductor will resist. But if you crank up the magnetic field suitably high then the superconductor will relent by allowing vortices to penetrate. These vortices repel, and so form a crystal-like structure known as an *Abikosov lattice*.

- The case $\xi = \sqrt{2}\lambda$ is of less relevance physically, because you have to fine tune two length scales, but is the situation with the richest mathematical structure. Now the attractive scalar force and repulsive magnetic force cancel, at least to leading order. Somewhat miraculously, it can be shown that this cancellation persists to all orders and the equations of motion exhibit solutions where k vortices can sit at k arbitrary points on the plane. These are known as *BPS vortices*.

Magnetic Monopoles are Confined

There is a lesson to take from the theory of superconductivity that will be important for particle physics. For this, we set up a thought experiment.

Our thought experiment involves a hypothetical object called a *magnetic monopole*, a particle that emits a radial magnetic field

$$\mathbf{B} = \frac{g\hat{\mathbf{r}}}{4\pi r^2}\ . \tag{2.118}$$

Here g is the *magnetic charge*. If you've been told that magnetic monopoles can't exist because the Maxwell equation $\nabla \cdot \mathbf{B} = 0$ is sacrosanct, then you've been lied to. (See,

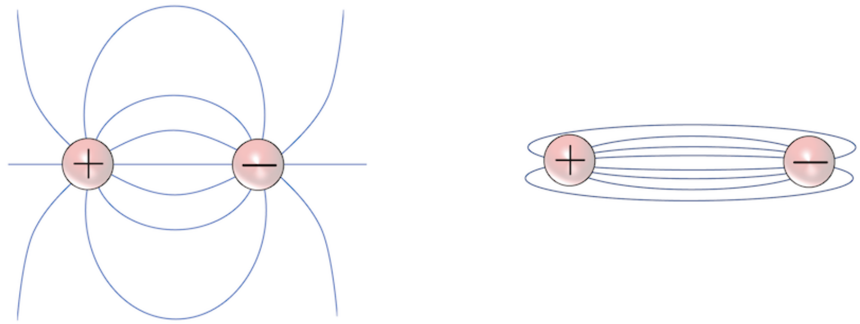


Figure 7. The magnetic field lines between a monopole anti-monopole pair. In a vacuum, the field lines spread out as a dipole configuration as shown on the left. But in a superconductor, the field lines form a flux tube as shown on the right, resulting in the confinement of magnetic monopoles.

for example, the lectures on Gauge Theory for a discussion of how magnetic monopoles are compatible with everything you know and love.)

Suppose that we have two magnetic monopoles, one with charge $g = 1$ and the other an anti-monopole with charge $g = -1$. If we place these monopoles a distance r apart in the vacuum, then the magnetic field lines will form the kind of dipole configuration that is familiar from our first course on Electromagnetism. This is shown on the left in Figure 7. The potential energy $V(r)$ between two monopoles scales like the Coulomb force,

$$V(r) \sim \frac{g^2}{r} \ . \tag{2.119}$$

Things are more interesting if we put the monopoles inside a superconductor. Now, the Meissner effect means that it's no longer energetically preferable for the magnetic field lines to spread out all over space. Instead, the field lines will clump together to form a magnetic flux tube that, at least far from the monopoles, is described by the vortex solution that we met above. A cartoon of the field lines is shown on the right of Figure 7. Now the potential energy scales linearly with the seperation,

$$V(r) \sim \mathcal{E} r \ . \tag{2.120}$$

where $\mathcal{E}$ is the energy per unit length of the vortex. This makes it very difficult to separate the monopole and anti-monopole: the further you want to pull them apart, the more energy it will cost. This is because they are attached by the flux tube which acts a little like an elastic band. (A little like an elastic band, but not a lot. Hooke's law is $V \sim r^2$ while here we have linear potential energy, $V \sim r$, corresponding to a constant force.)

Particles that experience a linear potential, like (2.120), are said to be *confined*. In Section 3, we will see that quarks in QCD exhibit a similar behaviour, albeit for more mysterious reasons.

2.3.3 Non-Abelian Higgs Mechanism

The idea of the Higgs mechanism extends naturally to non-Abelian theories. This is the context in which we will need it when discussing electroweak theory in Section 5.

One novelty is that the gauge group G need not be broken completely, and there could be some surviving massless gauge bosons. We will illustrate this with an example. Consider again the $O(3)$ sigma model that we previously discussed in Section 2.2 in the context of spontaneous symmetry breaking of global symmetries. This time, however, we will promote the $SO(3)$ symmetry to a gauge symmetry.

We have a 3-vector of real scalars, ϕ^a with $a = 1, 2, 3$ and define the covariant derivative

$$\mathcal{D}_\mu \phi^a = \partial_\mu \phi^a + g\epsilon^{abc} A^b_\mu \phi^c \ . \tag{2.121}$$

Here the ϵ symbol appears in its role as the generators for $SO(3)$,

$$T^a_{bc} = -i\epsilon^{abc} \ . \tag{2.122}$$

Alternatively, we could view this as an $SU(2)$ gauge theory with the field ϕ transforming in the adjoint representation. We consider the action

$$S = \int d^4x \ \left(-\frac{1}{4} F^a_{\mu\nu} F^{a\,\mu\nu} + \frac{1}{2}\mathcal{D}\phi^a \mathcal{D}^\mu \phi^a - \frac{\lambda}{2}(\phi^a\phi^a - v^2)^2 \right) \ . \tag{2.123}$$

Here $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\epsilon^{abc} A^b_\mu A^c_\nu$. In contrast to our previous Yang-Mills action (1.91), we've written the action in terms of the components of the gauge field, A^a_μ with $a = 1, 2, 3$ rather than packaging them into a 3×3 matrix. (This presentation turns out to be marginally simpler for the case of $SO(3)$.)

In the ground state, we have $\phi \cdot \phi = v^2$. We can make a choice of vacuum, say $\phi = (0, 0, v)$. When we were talking about global symmetries, we saw that this broke $G = SO(3) \to H = U(1)$ (or, equivalently, $O(2)$), and the same is true now the symmetries are gauged. This means that we expect a massless photon to remain, corresponding to $H = U(1)$, while the other two gauge bosons should become massive due to the Higgs mechanism. We will now see that this is indeed what happens.

As in the Abelian case, we sit in our chosen vacuum and look at fluctuations. The key is in finding the right parameterisation. We choose

$$\phi^a(x) = e^{i(\xi^1(x)T^1+\xi^2(x)T^2)}\begin{pmatrix} 0 \\ 0 \\ v+\sigma(x) \end{pmatrix} \tag{2.124}$$

with T^1 and T^2 the appropriate $SO(3)$ generators (2.122). If we were dealing with a global $G = SO(3)$ symmetry, then the fields $\xi^1(x)$ and $\xi^2(x)$ would be the Goldstone bosons. (They are related to the scalars that we called $\theta(x)$ and $\varphi(x)$ in the $O(3)$ sigma-model (2.56).)

Crucially, however, we're now thinking about the situation in which $SO(3)$ is gauged, and the two would-be Goldstones $\xi^1(x)$ and $\xi^2(x)$ can both be removed by an $SO(3)$ gauge transformation which acts on the scalar as $\phi \to e^{i\alpha^a T^a}\phi$ for some choice of $\alpha^i(x)$. In this way, they get eaten by the gauge fields A^1_μ and A^2_μ, just as in the Abelian case. In the resulting unitary gauge, the gauge fields and remaining fluctuating scalar $\sigma(x)$ are then described by the action

$$S = \int d^4x \, \left(-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma + \frac{1}{2}g^2(v+\sigma)^2(A^1_\mu A^{1\,\mu} + A^2_\mu A^{2\,\mu}) - V(\sigma)\right) \ .$$

with

$$V(\sigma) = \frac{\lambda}{2}\sigma^2(\sigma+2v)^2 \ . \tag{2.125}$$

As we anticipated, we have two massive gauge bosons, A^1_μ and A^2_μ, each with mass $m^2_\gamma = g^2v^2$. But the gauge boson A^3_μ remains massless. This is the photon associated to the unbroken symmetry group $H = U(1)$. There is also the massive Higgs field σ with mass $m^2_\sigma = 4\lambda v^2$.

As we commented previously, the gauge boson and Higgs boson have parameterically different masses, so it naively looks like it's possible to take a limit such that $m_\sigma/m_\gamma \to \infty$ and so we can decouple the Higgs and be left with a theory of only massive interacting gauge bosons. This time, however, the limit turns out to be problematic. This can't be seen in the classical analysis that we're focussing on here, but needs us to look more closely at the quantum amplitudes. Ultimately, it boils down to the fact that the theory of purely Goldstone modes is an interacting sigma-model (2.56) and, as such is non-renormalisable. This contrasts with the Abelian situation where the Goldstone that gets eaten is free before gauging. We will return to this issue in Section 5 when we discuss the Higgs mechanism in the Standard Model.